TONGSU SHUXUE MINGZHU YICONG

通俗数学名著译丛

DANGDAI SHUXUE

WEILE RENLEI XINZHI DE RONGYAO

让·迪厄多内 著

沈永欢 译

上海教育出版社

当代数学

为了人类心智的荣耀

当代数学

为了人类心智的荣耀

让·迪厄多内著　沈永欢 译　·　上海教育出版社





Jean Dieudonné

Pour l'honneur de l'esprit humain

Les mathématiques aujourd'hui

Hachette

根据阿歇特出版社 1987 年第 1 版译出，

本书中文版权由上海市版权代理公司帮助取得

通俗数学名著译丛

当代数学：为了人类心智的荣耀

［法］让·迪厄多内 著

沈永欢 译

上海教育出版社出版发行

（上海永福路 123 号）

（邮政编码：200031）

各地新华书店经销 商务印书馆上海印刷股份有限公司印刷

开本 850×1156 1/32 印张 11.25 插页 4 字数 267,000

1999 年 7 月第 1 版 1999 年 7 月第 1 次印刷

印数 1～5,100 本

ISBN 7-5320-6306-2/G・6461 定价（软精）：14.70 元

译丛序言

数学，这门古老而又常新的科学，正阔步迈向21世纪.

回顾即将过去的世纪，数学科学的巨大发展，比以往任何时代都更牢固地确立了它作为整个科学技术的基础的地位.数学正突破传统的应用范围向几乎所有的人类知识领域渗透，并越来越直接地为人类物质生产与日常生活作出贡献.同时，数学作为一种文化，已成为人类文明进步的标志.因此，对于当今社会每一个有文化的人士而言，不论他从事何种职业，都需要学习数学，了解数学和运用数学.现代社会对数学的这种需要，在未来的世纪中无疑将更加与日俱增.

另一方面，20世纪数学思想的深刻变革，已将这门科学的核心部分引向高度抽象化的道路.面对各种深奥的数学理论和复杂的数学方法，门外汉往往只好望而却步.这样，提高数学的可接受度，就成为一种当务之急.尤其是当世纪转折之际，世界各国都十分重视并大力加强数学的普及工作，国际数学联盟(IMU)还专门将2000年定为**"世界数学年"**，其主要宗旨就是"使数学及其对世界的意义被社会所了解，特别是被普通公众所了解".

一般说来，一个国家数学普及的程度与该国数学发展的水平相应并且是数学水平提高的基础.随着中国现代数学研究与教育的长足进步，数学普及工作在我国也受到重视.早在60年代，华罗庚、吴文俊等一批数学家亲自动手撰写的数学通俗读

物,激发了一代青少年学习数学的兴趣,影响绵延至今.改革开放以来,我国数学界对传播现代数学又作出了新的努力.但总体来说,我国的数学普及工作与发达国家相比尚有差距.我国数学要在下世纪初率先赶超世界先进水平,数学普及与传播方面的赶超乃是一个重要的环节和迫切的任务.为此,借鉴外国的先进经验是必不可少的.

《通俗数学名著译丛》的编辑出版,正是要通过翻译、引进国外优秀数学科普读物,推动国内的数学普及与传播工作,为我国数学赶超世界先进水平的跨世纪工程贡献力量.丛书的选题计划,是出版社与编委会在对国外数学科普读物广泛调研的基础上讨论确定的.所选著述,基本上都是在国外已广为流传、受到公众好评的佳作.它们在内容上包括了不同的种类,有的深入浅出介绍当代数学的重大成就与应用;有的循循善诱启迪数学思维与发现技巧;有的富于哲理阐释数学与自然或其他科学的联系;……等等,试图为人们提供全新的观察视角,以窥探现代数学的发展概貌,领略数学文化的丰富多采.

丛书的读者对象,力求定位于尽可能广泛的范围,为此丛书中适当纳入了不同层次的作品,以使包括大、中学生;大、中学教师;研究生;一般科技工作者等在内的广大读者都能开卷受益.即使是对于专业数学工作者,本丛书的部分作品也是值得一读的.现代数学是一株分支众多的大树,一个数学家对于他所研究的专业以外的领域,也往往深有隔行如隔山之感,也需要涉猎其他分支的进展,了解数学不同分支的联系.

需要指出的是,由于种种原因,近年来国内科技译著尤其是科普译著的出版并不景气,有关选题逐年减少,品种数量不断下降.在这样的情况下,上海教育出版社以迎接2000世界数学年为契机,按照国际版权公约,不惜耗资购买版权,组织翻译出版这套《通俗数学名著译丛》,这无疑是值得称道和支持的举措.参加本丛书翻译的专家学者们,自愿抽出宝贵的时间来进行这类

通常不被算作成果但却能帮助公众了解和欣赏数学成果的有益工作,同样也是值得肯定与提倡的.

像这样集中地翻译、引进数学科普读物,在国内还不多见.我们热切希望广大数学工作者和科普工作者来关心、扶植这项工作,使《通俗数学名著译丛》出版成功.

让我们举手迎接2000世界数学年,让公众了解、喜爱数学,让数学走进千家万户!

《通俗数学名著译丛》编委会

1997年8月

译者前言

当代数学是一个结构复杂、领域众多、宏伟壮丽的科学大厦.绝大多数数学家只能在其中一个领域甚至一个领域的某一分支毕生耕耘,能涉猎几个领域的数学家为数很少.这样他们就难于透彻地掌握当代数学的全貌及其来龙去脉.这或许可以部分地解释为什么几乎没有向广大受过教育的公众介绍当代数学的优秀通俗著作.只有在现代数学的多个领域做过深入研究、有过开创性贡献,并且通晓数学的历史发展,具有深邃洞察力和广博学识的数学家,才能写出这样的论著.现在我们有幸有了这样的书,就是让·迪厄多内写的《当代数学:为了人类心智的荣耀》.

迪厄多内于1906年7月1日生于法国里尔.在里尔和巴黎上了中学后,于1924年同时为综合工科学校和高等师范学校录取.他选择了当时处于巅峰状态的高等师范学校,在那里结识了许多日后成为大数学家的人物.1930年至1931年间,他得到洛克菲勒基金的资助,游学柏林和苏黎世.1931年在P·蒙泰尔指导下完成博士论文.1934年,他与A·韦伊,C·谢瓦莱,J·德沙特,H·嘉当等建立著名的布尔巴基小组.布尔巴基是对20世纪的数学进展具有广泛深远影响的学派,它的成员都是视野广阔,知识渊博,见解深刻,能够洞察数学理论的演变和本质,对基础数学的统一性具有坚定的信念,热衷于对基础数学给出尽可能完美的阐述和诠释.A·韦伊和H·嘉当都得过沃尔夫数学奖,A·韦伊还被美国数学家P·R·哈尔莫斯赞为"当代最伟大的数学家

(注意,不是'最伟大的数学家之一'——引用者)的候选人".1939年,布尔巴基学派开始出版多卷巨著《数学原理》,已出40多卷.

迪厄多内曾先后在法国、美国、巴西的多所大学任教.他发表过近150篇研究论文.迪厄多内在本书中把当代数学划分为21个领域(第五章§5,A),而据译者粗略分析统计,其中他本人有过独创性研究的领域,竟达到12个,占到一半以上!迪厄多内具有超人的效率,写有关于线性代数和初等几何、微积分、典型群的几何学、代数几何学、形式群理论、经典群的自同构等多本专著(其中《典型群的几何学》有万哲先的中译本).特别应当指出,从1960年至1982年,出版了他写的9卷本《数学分析原理》,这是关于现代分析的系统辉煌论述,有英文、德文、俄文译本,其中第一卷有郭瑞芝、苏维宜的中译本,第二卷有拙译的中译本(《现代分析基础》),科学出版社,第一卷,1982;第二卷,1986).再有,在布尔巴基前30卷著作中,他写的篇幅超过其他几个人写的总和.

迪厄多内对数学史有浓厚的兴趣和深刻的研究.他著有关于泛函分析史、代数几何学史、代数拓扑学和微分拓扑学史以及关于现代数学概观(布尔巴基的看法)的论著,主编了《1700—1900年数学史纲要》.此外,布尔巴基关于数学发展的历史注记,大都出自他的手笔.

迪厄多内于1968年当选为巴黎科学院院士,1978年被授予荣誉勋位勋章.他殁于1992年11月29日.逝世次日,巴黎科学院终身秘书P·热尔曼写道,J·迪厄多内是"我们时代数学的巨人,更确切地说,他熔铸于当代数学中,他就是当代数学".《世界报》于1992年12月2日发表的文章中称他为"当代数学的象征",J·-P·皮埃认为他是"当代数学最杰出的人物之一",是"在数学中具有百科全书式才智的数学家".

从上面的简述中不难看出,迪厄多内一直在现代数学主流

中搏击，对其发展和全貌有着完整深刻的了解.以他的水平和功力写一本关于当代数学的通俗论著，一定会不同凡响.果如其然，本书出版后，很快就出现了多种文字的译本.据译者所知，1989 至 1990 两年内，就出现了日文、西班牙文、葡萄牙文、意大利文和英文译本.迪厄多内的著作通常都有几种外文译本，但本书外文译本语种之多，在其著作中首屈一指.

本书是为广大受过教育而又对科学尤其是数学感到兴趣的公众写的，因此作者限于从代数、数论和集合论中撷取例证.作者在书中着重阐明数学在现代其实经历了真正的变革.如果说 19 世纪以前数学的特征之一是具有高度的抽象性，那么现代数学则更加抽象，它研究的是数学结构，其主要特征是研究对象之间的关系而不是这些对象本身的具体性质，因此它更加得不到外在的、可以感知的"形象"来显现或支撑.但是，这种变革又是必然的、自然的.为攻克经典时代遗留下来的数学问题或其他科学部门要求数学解决的问题，数学家们必须创造成为当代数学发展主流的对象和方法.

为了阐明作者力图解释的重点内容，迪厄多内首先简短而又生动有趣地谈论了数学的概念，数学家和数学界的特点以及数学问题的性质.接着作者从论述经典数学的对象、方法和几种不同类型的问题开始，勾勒当代新的数学对象和方法是怎样产生的，它们具有什么样的基本特征，以及当代数学总的面貌如何.最后作者还讨论了数学基础问题.

作者把需要较多数学知识(理工科大学一二年级水平)的内容写进各章的附录中，因此，为理解本书主要章节，具有高中数学水平也已足够.当然，由于数学总究是一门推理严谨的科学，因此需要读者在阅读大部分章节时集中精力，勤于思考.

译者以为，本书对专业数学工作者和广大数学教师也极有教益.译者就从翻译本书中获益匪浅.当然，限于译者水平，译文中舛错恐或难免，尚祈读者不吝指正.

“傅里叶先生认为，数学的主要目的是服务人类、解释自然现象；但像他这样的哲学家应当知道，科学的唯一目的是为了人类心智的荣耀，因此，一个关于数的问题与一个关于宇宙体系的问题具有同样的意义.”

——C·G·J·雅可比

1830年7月2日致勒让德的信，

见其 Gesammelte Werke(全集)，

第1卷，柏林(Reimer)，1881，PP.454.

献给

奥代特和弗朗索瓦



目　录

附录

导　言

本书是特地为这样一些读者写的：他们由于各种原因对科学感兴趣，但不是职业数学家．经验表明，情况大致总是这样：虽然这些人喜爱阅读和听取关于自然科学的讲解，并且感到从这些讲解中获得了知识，开阔了眼界，但他们发现关于当代数学的文章都是用无法理解的行话写就，而且讨论的概念过于抽象，使人趣味索然．

本书的目的是试图解释这种对数学缺乏理解的现象的原因，或许还想打破这种隔阂．由于总得认定人们必须具有一些最低程度的数学知识，否则就不可能一起讨论数学，所以我得假定本书读者已经学过相当于科学业士学位[①]水平所要求的课程．我以各章附录的形式插进一些增补内容，这是为多少学过一点科学课程（至少达到大学二年级结束时的水平）的读者准备的；但为理解正文，没有必要去读这些附录的任何部分，因此略去各章附录不会有什么损害．所必须的只是要做一点努力，在领会一连串推理（其中互相连结的每一步本身是完全初等的）过程中能集中注意，专心致志．

---

① 法国的业士学位考试即中学毕业会考，对中学毕业生（通常为19岁）进行，以决定他们能否进入大学．就中国情形而言，学过相当于大学理工科一年级高等数学课程的读者只要认真思考，即能读完本书；对数学有兴趣的高中毕业生也能读懂本书一大部分．——译注

我打算揭示的是,这种"最低水平"的本性,正是人们理解当代数学过程中经历的困难的根源.中学里所教的数学,没有什么东西是1800年以后发现的[①].对于自然科学中除物理学以外的所有分支中认为必须具备的数学知识,情形并无很大差别.即使在物理学家中,我相信除研究量子理论或相对论的人之外,那些从事实验工作的人几乎不会用到比麦克斯韦在1860年时所知道的更多的数学知识.

[9] 人人都知道,自从19世纪初以来,自然科学以异乎寻常的方式不断发展.受过教育的公众所以能跟上这种令人头晕目眩的进展,当然要落后几步,但还不至于迷失道路,是得益于保留了新思想实质的巧妙简化.数学恰恰也以同样的方式前进,但除数学家外,很少有人意识到这个事实.其原因在于,一方面,如前所说,大部分科学即使今天仍然很少需要超过"经典"数学[②]范围的任何东西;另一方面,在数学的演进中发生了真正的突变,与"经典"数学对象——数和"形"——相当不同的新的"数学对象"创造了出来.与"经典对象"相比,这些对象具有严格得多的抽象性(它们在任何方面都得不到看得见的"形象"的支持),这就吓退了从这些对象身上看不到任何用途的人.

我愿意使对于数学有好感的读者相信,这种抽象决不是来自数学家的反常意愿,似乎他们想通过使用深奥莫测的语言来把自己与科学界隔开.数学家的任务是求出"经典"时代传下来的或直接来自物理学中新发现的问题的解.他们发现这有可能办到,但只能通过创造新的对象和新的方法,而这些对象和方法的抽象特征对于它们的成功是必不可少的.

因此,我试图表明,1800至1930年间的数学家是被经典对

① 但不是1800年以前发现的一切都已在中学里得到讲授,远非如此!

② 我用"经典数学"来指1800年前所知道的全部数学结果.

象和关系的本质特性逼着(常常那时并不认识到这一点)去锻造新的"抽象"工具,使得他们及其当代继承者能够解决过去看来是不可攻克的问题.当代数学家能富于想象力地使用这些工具并扩展它们的范围,但与有些时候所想的相反,他们并没有创造这些工具.

我认为,对过去数学的主要特征没有某些最低程度的了解,就不可能理解当代数学.本书最初两章概述数学和数学家在当今世界中的地位,以及数学与其他科学的联系.接着的第三章则是对于从欧几里得到高斯的经典数学的简略概观.这一章的重点放在古希腊人的伟大独创思想上,就是以一种不可磨灭的方式赋予数学对象以"思维对象"的特征.第四章给出一些例子,它们是从经典数学中出现的成千上万个问题中挑选出来的,这些问题在 19 世纪引起过重大研究. [10]

第五章是我所作论证的核心部分.我们还是通过几个读者能够理解的例子来揭示,在 19 世纪中,怎样从分析上世纪遗留下来的各种问题逐渐导致发现这些问题的真正性质,从而铺平了部分或全部解决它们的道路.但为此必须付出放弃经典数学对象所具有的半"具体"特征这样的代价;应当了解,关于这些对象,本质的东西不是它们表面上具有的个别特性,而是它们之间的关系.常有这样的情形,对于外表非常不同的对象,这些关系却相同,因此必须以扬弃外表的方式来表述这些关系;例如,如果我们想规定一种关系,它既可定义于数之间,也可定义于函数之间,那么它只能通过引进这样的对象才能办到:它既不是数,也不是函数,但又可以使之特定化为数或函数,或其他类型的数学对象.正是这些最终得到研究的"抽象"对象被称为数学结构,在第五章中我们会描述一些最简单的结构.

在 19 世纪,数学家们试图找到一条道路,它可以完全精确地表述定义和定理,使任何不确定性无容身之地;在最后的第六章中,我们要讲述这样做的过程中出现的困难.为此,一旦欧几

里得《几何原本》的不精密性得到改正,而且古希腊人独创概念的纯粹性得到重新发现,《几何原本》就被用作数学所有分支的样板;在伴随数学应用于经验科学所取得的令人难以置信的成功①而来的欢欣快慰中,有些事情过于长久地被忽略了.

喜欢引起轰动的图书的人会对本书感到不满,他们在这里找不到震撼世界的充满机智反论的论点.我把自己限制在陈述有关数学的事实,而不是表达看法上,而且我注意不去卷入争论.我同样注意约束自己,不试图对我所讨论的概念的未来演变作任何预测.

可进一步阅读的文献

[1] J.Dieudonné 及其 10 位合著者,Abrégé d'histoire des mathématiques, [11] 1700—1900,新版,Paris(Hermann),1986.

[2] G.Choquet,La naissance de la théorie des capacités: réflexion sur une expérience personnelle,"La vie des Sciences",C.R.Acad.Sci.,série générale,3(1986),PP.385 - 397.

[3] P.Dedron,J.Itard,Mathematics and mathematicians,5 vol.,J.V. Field 译自法文,Milton Keynes(Open University Press),1974.

[4] Diophantus of Alexandria,Les six Livres d'arithmétique et le livre des nombres polygones,由 P. Ver Eecke 翻译,Bruges(Desclées de Brollwer),1926.

[5] Euclid,The thirteen books of Euclid's Elements,3 vol.,T.Heath 翻译,New York(Dover),1956(中译本:欧几里得,几何原本,陕西科学技术出版社,1990).

[6] L.Euler,Correspondance avec C.Clairaut,J.d'Alembert et J.-L. Lagrange,Opera Omnia,ser.quarta A,Basle(Birkhäuser),1980.

[7] G.H.Hardy,A mathematician's apology,第一版,Cambridge(Cam-

① 正如 E·威格纳所说,对于构造适合经验世界的模型,甚至物理学家也对"数学的超乎常理的灵验"感到惊异.

bridge University Press),1940(中文节译本:G.H.哈代,一个数学家的自白,数学译林 3(1984),PP.267-273,383-388).

[8] S.Lang,The beauty of doing Mathematics,New York,Berlin,Heidelberg-Tokyo(Springer),1985.

[9] H.Poincaré,a)La Science et l'hypothèse,Paris(Flammarion),1902(中译本:H·彭加勒,科学与假设,商务印书馆,1989);b)La valeur de la science,Paris(Flammarion),1905;c)Science et méthode,Paris(Flammarion),1908(a),b),c)合一中译本:H·庞加莱,科学的价值,光明日报出版社,1988);d)Dernières pensées,Paris(Flammarion),1913.

[10] A.Weil,Oeuvres scientifiques,3 vol.,New York-Berlin-Heidelberg-Tokyo(Springer),1979.

由于本书不是为专家学者写的,所以我不想让正文充塞大量的"参考文献",而且我也只限于举出几本其精神实质与本书相一致的论著.我向读者保证,对于我所作的涉及数学史的每个论点,我都能征引权威出处;我希望读者相信这句话.

在鼓励读者获知数学史的更多详情方面做得太多是不可能的.我想请读者注意论述[3].该书由于有很多插图并且不需要任何超过业士水平的知识,因而很有吸引力.对于水平更高的读者(大学二、三年级水平),请允许我指出论述 1700 年后数学史的著作[1].

研读丢番图[4]和欧几里得[5]总是富有教益的;T·希斯翻译的《几何原本》英译本中有大量数学或哲学性的注解.

数学家中基于个人体会发表关于他们所研究主题的沉思或反省的人并不多.欧拉的大批信件[6]使我们得以更好地了解他 [12] 的生活和工作方法;而读读庞加莱关于数学所写的许多文章(收集在[9]中开列的4本书中),总会是有趣味的.离我们更近时代的数学家 G·H·哈代在晚年写了极有感染力的"自白"[7],书中常显出作者不在乎自己数学家的身份.

我们时代最伟大的数学家之一 A·韦伊想出了一个巧妙的

主意,就是对他的选集[10]中的每本书或文章加上内容广泛的评注,这相当于一本激动人心的"工作日记",几乎逐日勾勒了一个富有创造性的数学家的思想演变脉络.尽管充分鉴赏这些内容需要极其广阔的数学文化背景,但由于它们既非自吹自擂,又不故作谦逊,所以不会有人不能从这些回忆中获得教益.

完全类似而且内容同样丰富的是G·绍凯最近写的[2],书中描述了通向他最优秀的发现之一——广义容量——的道路.

最后我要热切地指出法裔美籍数学家S·兰最近的壮举,他在探索馆的三次会议和讨论上,把一批非同寻常的公众引向理解"数学之美"([8]).演讲人——尽管他是一位"纯粹"数学家——甘心情愿把自己置于听众支配之下,以及他在面对不可预料的听众反应时所表现的镇定自若,都使此书成为既生动有[13]趣又富于教益的小品.

第一章
数学与数学家

1. 数学的概念

数学在人类活动中所占的地位是自相矛盾的.当今"发达"国家中几乎每个人都知道,数学是一门重要学科,大多数科学技术分支都用得着它,而且比以往任何时候都要多的职业,如果没有某些数学知识,就无法从事.然而,如果你问"数学是什么?"或问"数学家做些什么?",那么除了荒唐可笑的回答外,你几乎会一无所获,除非你的对话者至少经历过大学二年级的数学训练.即使别的科学领域里的著名专家,对于数学家的工作也往往只有一些反常的看法.

由于每个人都是在上小学时通过数字演算接触数学的,所以传播最广的看法是,数学家就是从事这些演算的内行里手.随着电子计算机及其语言的出现,人们现在会认为这意味着数学家就是特别善于为这些演算"编制程序"的人,他们整天做的就是这件事情.工程师们老是在寻找他们感兴趣的量的最优值,他们把数学家看作储藏公式的仓库保管员,应他们的要求为他们提供所需的公式.但是眼下由于媒体大量报导别的科学部门所取得的进展,因此几乎所有当代人都坚定地持有这样一种最能导致误解的看法,即数学学科已不再有什么东西可以去发现,数学家们的工作只能限于把以往年代的遗产传下去.

我们可以对这种状况高枕无忧,而且除了同意A·韦伊所

说,"数学之古怪在于它不能为非数学家理解"([10],Ⅱ,PP.
[15] 465),再也不做更多的事情吗?我想我们至少应当试着找出这种缺乏理解的原因.事实上,有许多期刊致力于向大量读者以各种科学教育水平普及新近的科学发现.除了我们将会提到的少数例外(第六章,§5,C),关于数学中的最近进展,没有任何这类介绍(它会使人觉得数学中什么也没有发生!);与此相反,却有很多关于天体物理学、地质学、化学、分子生物学甚至原子或核子物理学的通俗文献.看来借助略去细节并使实质得以清晰显示的图表,再加把专家从复杂细致的实验中所能导出的结论予以根本简化的阐释,就有可能使这些刊物的读者觉得,他们已经懂得原子、基因或银河是什么,尽管这些东西的数量级从埃(10^{-10}米)延伸到光年(3.10^{15}米),而且超出我们想象所能掌握的范围.

相反,让我们提出现代数学中最富有成果的理论之一即称之为"层的上同调"的理论.它的草创是在1946年,多少与分子生物学中的"双螺旋"结构处于同一时期,并且也取得了可与后者媲美的巨大进步.然而,我却不大可能对没有读过大学前两年数学课程的人解释这一理论的内容.即使对具有这样水平的好学生,我的解释也得花上几个小时;至于阐述这种理论怎样应用,那得花多得多的时间.这是因为这里我们不再能用解释性的图表;在我们能接触到所讨论的理论之前,我们必须消化十来个同样抽象的概念:拓扑、环、模、同态,等等,这些概念中没有一个能用"形象化"的方式来表达.

对构成当代伟大数学理论的所有观念,上面说的话几乎都适用;第五章§5,A中列举了这些观念,读者可以查阅.因此我必须甘心情愿地接受这一点:如果我想在业士所能接受的水平上进行讨论,那么我只能在第五章§5,A中列举的21个论题中提及其中3个即代数、数论和集合论的基本概念,不时对知道一点微积分的读者来点评注.我将试着从这些概念开始,解释它们

如何由必然性驱使而逐步发展,显示出蕴藏在其中的抽象得多的概念,在此过程中获得了解决数学问题的不可比拟的巨大效验.再向前进就意味着把词汇的溪流汇集起来,这样就会使不是 [16] 从事数学专业的读者很快迷失方向.我所提到的那些内容,只是使读者得到我所讨论的一些定理的基本想法,它们与现代数学知识的巨大容量相比,仅仅是极其微小的一部分.

2. 数学家的生活

通常"音乐家"这个词可以指作曲家、演奏家、音乐教师或者兼作其中几项的人.同样,说到"数学家",我们可以理解为数学教师,应用数学的人,或者创造性的数学家;而普遍的看法是最后这类人已经绝种.可是,由于我的目的是帮助读者理解现代数学的起源和性质(我们将会看到[第五章 §5,A],现代数学的大部分内容,都是 1840 年以后建立起来的),因此我们讨论的几乎都是最后这类人即创造性数学家.这样,下面谈到的数学家,就定义为至少发表过一条非平凡定理的证明的人(稍后在 §4 中我们会精确定义"平凡"指什么).

在我看来,在我们试着解释数学家做些什么之前,让读者到数学家们的陌生国度小小游历一番,不会是毫无趣味的.

有充分理由认为数学创造能力与种族无关.从整体上看,似乎这种能力也不能遗传:数学家的后代成为数学家总是少有的例外.但是社会环境对数学天才的发育起着重大作用,即使在最伟大的天才中,也没有数学家从头开始就自己重新发现他那时代的数学知识这样的例子,与此有关的关于帕斯卡的故事不过是一则传说.社会环境必须能使未来的数学家至少接受一些初等教育,向他显示名副其实的证明,唤醒他的好奇心,然后使他能接触到他那时代的数学,即得以按照自己的条件读到阐述性的论著.直到 18 世纪末以前,还没有真正有组织的大学数学教育;在此之前,几乎所有伟大的数学家,从笛卡儿和费马到高斯

和狄里克雷,都是通过研读他们杰出前辈的论著(这还是非常好的训练!)而自学成材的.即使在今天,很可能许多数学天才由于[17]缺乏有利的社会环境而不能显现;我们不知道欠"发达"社会中的数学家,也就不足为奇.即使在更为发展的国家,当初等教育受到宗教或政治的限制或只有功利主义的考虑时,它也会不利于数学才能的养育,就像已经进入20世纪时美国的情形那样.那里中等教育实际上掌握在地方当局手中,他们常常习惯于认为,对于青少年学会打字或开车比学习拉丁文或欧几里得几何学更加重要.我认识一些著名的美国数学家,他们生在小镇上,18岁以前从未见到过数学证明;由于被科学吸引,他们在大学里接受面向工程职业的训练,只是出于偶然的机会他们才接触到象样的数学教育,由此发现了他们真正的才能.

在充分广泛的社会阶层能够获得教育机会(必要时借助助学金)的社会里,数学家的出身是多种多样的.他们可能来自名门望族,如法格纳诺、黎卡提和达朗贝尔或来自有产阶级上层,如笛卡儿、费马、帕斯卡、克罗内克、若尔当、庞加莱和冯·诺伊曼,他们也可能来自卑微的下层,如罗伯瓦尔、高斯、埃利·嘉当和勒贝格.大多数数学家生于中等家庭,也常有生于贫困之家的.

通常数学才能总在15岁前后被发现,但也可能为没有任何证明概念的教育延后,如上面说到的美国的情形那样.无论如何,与传播相当广泛的看法相反,创造阶段极少在23至25岁前开始;帕斯卡、克莱罗、拉格朗日、高斯、伽罗瓦、闵科夫斯基以及当代的德利涅等是例外,他们在20岁前就作出了著名的数学发现.至于数学创造的寿命,另一个流行看法是——如哈代所说([7],PP.10),"数学是年青人的玩意".确实,很多重要发现是由比30岁大不了多少的数学家作出的;但仍有许多伟大数学家,如庞加莱、希尔伯特或H·外尔,其创造阶段延长不少,他们直到50或55岁还保持相当多产的状态.另外一些人,如基灵、

埃米·诺特、哈代(如他自己所说([7],PP.58))或晚近的扎里斯基和谢瓦莱,在40岁前后才做出他们最好的成果.最后,有一些为众神钟爱的数学家,如克罗内克、埃利·嘉当、西格尔、A·韦 [18] 伊、J·勒雷和I·盖尔范德,60岁后还在继续证明优美的定理.但是,就像许多体育运动中很少有人能指望在30岁后还占有冠军宝座,大多数数学家在55或60岁后由于看出自己创造性想象力的枯竭而不得不退下来.对于有些人,这是令人心碎的经历,如哈代在其"自白"中生动地描绘的那样;另有些人则通过干些他们还能做的活来适应.

像许多别的学者一样,数学家的生活充满着永不满足的好奇心和近于恋情似地解决他正在研究的问题的愿望,这能使他同周围的现实世界完全隔开.一些著名数学家心不在焉,行为古怪,其源盖出于此.事实是,数学证明的发现,通常总是在一段紧张持久的集中精力思索之后才能达到,有时在找到所求答案之前要成年累月地反复更新.高斯本人承认,他花了好几年去求一个代数表示式的正负号;庞加莱在有人问他如何作出其发现时答道:"经常对它们加以思考";他还详细地披露通向他的最好结果之一——发现"富克斯函数"的尝试和沉思过程([9]).

这样,数学家首先追求有充足的时间供他支配,使他能专心致志于自己的工作,这就是为什么19世纪以来他们都喜欢在大学或学院中任教,因为那里授课时间相对较少而假期又比较长.报酬只居次要地位;最近我们看到,在美国以及其他一些国家,数学家宁愿以工资下降不少为代价,放弃工业界中薪俸优厚的职位而回到大学中去.

再则,只是到晚近,大学教职数才变得相当地多.1940年以前大学数学教授职位极为有限(在法国不超过一百);直到1920年以前,像库默尔、魏尔斯特拉斯、格拉斯曼、基灵和蒙泰尔这样有水平的数学家不得不毕生或一段时间满足于呆在中等学校里教书;情况在很长时间里一直是,小国几乎没有什么大学.

在 19 世纪以前,数学家从事自己想做的工作的机会甚至更加不稳定.如果没有个人财产、赞助者或科学院提供还算不错的[19]生活来源,那么除了当天文学家或测量人员(高斯就是如此,他在这些职务上花了大量时间)之外,几乎就不再有别的出路.无庸置疑,正是这种情形解释了为什么 1800 年前富有才华的数学家数目极其有限.

需要长时间致力于思考他们试图解决的问题,这就几乎自动排除了在其他生活领域(例如行政管理)或严肃科学工作中同时完成需要全神贯注的任务的可能性.傅里叶在担任伊泽尔地区行政长官时创造了热的理论或许是绝无仅有的情形.至于数学家出任行政管理或政府显要职位,我们会发现他们在任职期间实际上放弃了他们所从事的研究.此外,正如哈代嘲讽地评论的那样([7],PP.12),"离开数学的数学家的最近记录不是特别令人鼓舞的".

至于数学家的个性,如果除去他们中许多人常常心不在焉和缺乏"实际技能"外,那可以说他们与随便哪个地位相仿的阶层并没有什么区别.在他们身上同样可以发现形形色色的性格以及各种各样美德和缺陷.如果他们有时会对自己具有的非同寻常的才干感到骄傲和自豪,那他们必须记住,许多别的技巧也只能由很少一部分人分享.无疑能在高空钢丝上跳舞的演员或能下盲棋的棋手还要少得多.然而,正如哈代所说([7],PP.16),数学家的行为"确实具有坚韧不拔的特性",而如果你在人类知识的金字塔上添加了一小块石头,那你就没有虚度年华.

无论如何,数学家必须承认,他们的才能并没有授予他们在数学之外领域中的任何特殊本事,而那些自认为由于在同事中享有威望就能去改造社会的数学家到头来尝到了苦果.

数学家的宗教倾向或政治立场更是多种多样:柯西是偏执的天主教信徒,而哈代则是古怪的无神论者,上帝对于他是个人的敌人;高斯非常保守,而伽罗瓦则是热情的革命家.一个极端

却又令人痛心的例子是年青的德国数学家O·泰希米勒,他或许具有可与伽罗瓦相媲美的天才,但却是狂热的纳粹分子,为驱逐他的犹太人师长出力.他作为党卫军军官赴苏德战线,一去不返.

大多数数学家过着好公民的生活,对震撼世界的动乱有点担心,不指望权力和财富,安于小康境地.他们不急切地追求荣 [20] 誉,但对它也并非毫不在乎.与许多艺术家相反,他们很少投入激情的漩涡中,在他们中间很难探到轰动一时的情侣或浪漫史,他们自己也只能满足于伽罗瓦、柯瓦列夫斯卡娅或拉马努金的多少带点浪漫色彩的传记.

3. 数学家的工作与数学界

实验性科学的研究是在实验室中进行的,实验室需要越来越大的队伍来操作仪器,仔细检查结果.做数学研究除了纸张和藏书丰富的图书馆,再也不需要别的什么.像实验性科学中实行的集体协力工作,在数学中相当罕见;许多数学家发现,除非在安静的、与世隔绝的环境中,否则就很难认真地思考问题.虽然合作是相当普遍的,但通常都是参与合作的每个人拿出自己独处时搞出的名堂,再把各自的结果集中起来,这样可以互相从对方的想法中获得启发,使他们能从新的起点进一步前进.数学家中长期合作的最著名例子是哈代与李特尔伍德,一位住在牛津,另一位住在剑桥,两人只是偶然谋面,他们的共同工作完全靠通信进行.

然而,如果说数学家的大多数出版物是个人的作品,那么像格拉斯曼、享泽尔或埃利·嘉当那样长期在几乎完全与世隔绝状态下还保持多产的数学家则为数极少;这三位是由于想法新奇,不能为同代人理解.大多数人如果缺乏经常与同行交流的机会,以便他们能被同行理解,就会感到沮丧和失望.由于他们的研究的抽象性质使得难于同非数学家交换想法,情况就更是如此.

直到 17 世纪中期,仅有的交流手段是私人通信(有时以乐于助人的写信者,如梅森和科林斯为中心)、相互访问以及有时由作者出资印书(除非有资助者同意付款).由学术社团定期出[21]版的第一个期刊大约出现于 1660 年,但为数一直很少,而且 1820 年前出版的为数很少的科学期刊都是而向一般科学的,并无专业分工.第一个跨越国界的专门数学期刊是创办于 1826 年的"克雷尔杂志"①,紧接着有 1835 年创办的"刘维尔杂志"②.

然而正是现今,语言障碍开始出现.在 17 世纪,几乎所有科学出版物都以拉丁文写就(笛卡儿的《几何学》以及帕斯卡的某些著作是仅有的值得注意的例外).这个传统在 18 世纪开始衰落,法国数学家那时已只用法文写作,法国以外的某些教学性著述则以作者的母语写就.高斯时代之后,科学书刊用拉丁文写作很快消失,而这就不断放慢丁新结果的传播速度,尤其在法国,那里直到 19 世纪末以前几乎不学习任何别的语言.颇为滑稽的一个例子出现于 1880 年,当时巴黎科学院为解决一个数论问题举行一次比赛,而该问题早于 20 年前已为 H·J·S·史密斯解决.显然埃尔米特和若尔当看不懂英文.众所周知,英文在当代合理地成为通用科学语言.

在 19 世纪,数学期刊的数目不断稳定地增长,1920 年后由于数学研究得到发展的国家增多,数学期刊的数目增加得更快,1950 年后更是臻于极盛,达到了爆炸的程度,以至现在世界上大约有 500 种数学期刊.为使人们不至于被这个信息大海淹没,

---

① 此刊物全名为"纯粹与应用数学杂志"(Journal für die Reine und angewandte Mathematik),由于创始人是 A·L·克雷尔,所以通常称为"克雷尔杂志".——译注

② 此刊物全名为"纯粹与应用数学杂志"(Journal de Mathèmatiques Pures et Appliquées),由于创始人是 J·刘维尔,所以通常称为"刘维尔杂志".——译注

从19世纪后30年起，创办了专门对其他出版物进行编目和摘要的期刊.这种期刊中传播最广的是美国的《数学评论》(Mathematical Reviews)，现在已几乎扩展到一年4 000页(每页平均有5到10条分析性条目).

随着期刊数目倍增，也出现了不少教学性的著作，它们常以丛书的形式组合起来(其中不少是相当专门化的).因此现在极少有一种新理论需要等待10年以上才能成为教学性论述的主题.

19世纪中叶，德国大学里出现了举办“讨论班”的活动：在一位或几位教授指导下，若干当地的或前来访问的数学家，其中常有攻读博士学位的研究生，在一个学年中定期举行会议，分析一个问题的现状或考察评述最值得注意的新结果.1920年以 [22] 来，这种活动推广到全世界，讨论班上所作的讲述常印成文章，以便能使更多人看到；在许多大学中为高水平学生开设的专门化课程也常常这样做.

然而，长期以来很明显的是，在科学领域中，口头交流往往比阅读论文有效得多.中世纪起，一个大学的学生到另一个大学游学就成为一种传统，它一直保留至今，尤其是在德国.再则，邀请学者到他们所在大学之外的别的学校去工作或讲学也已变得相当普遍.

从1897年的国际数学家大会起，这种个人接触的需要得以制度化了；1900年以来，每四年举行一次国际数学家大会(由于世界大战中断过两次).大会参加者数目不断增多，这在某种程度上削弱了它的功用；因而从1935年起，增加了各种较有限制的聚会：学术讨论会、专题会议、研究班、夏季讲习班等等，在这些聚会上，范围多少不太固定的专家相叙在一起，估量他们新近的发现，讨论他们当前面临的问题.

4. 大师和学派

竞争的心态，对于名次的追逐，为记录和奖金而比赛，这些

都从来没有像今天这样强烈.这在体育运动中尤为显眼,而同样的精神也扩展到了诸如象棋或桥牌这样的智力消遣,甚至扩展到许多别的多少有点可笑的领域.

尽管如此,我们近来看到这样一种观点得到热情的宣扬:所有人在智力创造性上都具有同样的能力,而我们在这方面所发现的不合理的差别只依赖于个人受惠于教育的多少.这是一种奇怪的学说,它需要大脑与别的器官相比具有不同的生理学这样的信念来支撑.当我们想起那些王公贵族或亿万富翁的子女,尽管有最好的家庭教师照料但仍蠢得不可救药时,对此我们只能耸耸肩膀.

我们必须承认,如同所有别的学科一样,在数学家中各人才能差别很大.在"发达"国家,数学教学是由一大批大学教师担当[23]的,他们中大多数通过进行一项独创性研究获得了博士(或与之相当的)学位,但决不能说他们在研究能力上全都相同,因为其中不少人可能缺乏创造性的想象力.

事实上绝大多数这种工作属于我们称为平凡的工作,就是说,它们只限于从熟知的原理中引出明显的结果.这些数学家的学位论文通常甚至都未发表;这些论文是在"导师"启发下写成的,更多地反映了后者的想法而不是作者的观点.因此,一旦离开了导师,这些人就很少发表与其学位论文直接有关的文章,随后很快停止发表有创造性的东西.尽管如此,他们仍具有无庸置疑的重要性:除了在国家科学精英的教育中他们所起的主导作用外,正是从他们中间在大学最初几年里就能挑出将会成为下一代数学家的出类拔萃的学生.如果他们能跟上自己所从事的学科的现代进展,并用这种知识丰富他们的教学,他们就能唤醒犹豫不决的有才华的学生,把这些学生推荐到能指导他们迈出作为未来研究者的最初几步的同事那里去.

在较高的水平上,我们发现人数少得多的一类(尤其在美国这样的国家,那里需要大量教学人员),就是能超越他们学位论

文中所做工作的界限，甚至涉及完全不同领域的数学家群；他们常能在研究中保持活跃达30年左右，而且发表几十篇独创性论文.在"发达"国家，把好的年份和差的年份算在一起，你可以说在一千万人中，每年会诞生一位这样的数学家；这样，像法国这样大小的国家大约有150位活跃的有创造性的数学家，而像美国和前苏联，则大约有600位.只有他们才能有效地承担研究生课程的教学工作，通过这些教学传播新的观念，也只有他们才能有效地指导年青数学家从事研究工作.

最后是伟大的革新家，他们的思想使他们时代的数学整个改变进程，其影响有时要超过一个世纪.但是，正如爱因斯坦对保尔·瓦莱里所说，"一个想法可真是珍贵呀！"(显然爱因斯坦在这里指的是伟大的想法).在18世纪，我们可以数出六、七位这样的天才人物，19世纪则可数出三十位上下，在当代我们可以指望全世界一两年中出现一位.同这样的伟大数学家一起工作，是一种使人精神振奋、获得丰富教益的经历；哈代说他同李特尔伍德和拉马努金的合作是"他生命中的决定性事件"([7],PP.88)，对于我同布尔巴基小组的合作，我也能说同样的话. [24]

诺贝尔奖中没有数学奖，但从1936年起，每四年举行一次的国际数学家大会每次开会时颁发两枚、三枚或四枚称为"菲尔兹奖"的奖章给其工作由一个国际委员会裁定为最杰出的数学家，得奖人最好是年龄小于40岁的.迄今为止共有10位菲尔兹奖得奖人来自美国，还有5位法国人，4位英国人，3位斯堪的纳维亚人，2位日本人，2位俄国人，1位德国人，1位比利时人，1位中国人，1位意大利人[①].总能批评这些奖的分发办法甚至分发原则，但我想没有人能对这些获奖者的份值提出怀疑，他们都

① 以上说的是1986年前的情形.在1990年和1994年两次大会上，又有8位数学家获得菲尔兹奖，其中有2位法国人，2位俄国人，1位美国人，1位日本人，1位比利时人，1位新西兰人.——译注

在其科学生涯中至少提出过一个伟大想法.此外,我们应当在这个名单外再添上 15 位左右毫不逊色的数学家,他们由于奖章数目有限而未获奖.

对于数学家,国籍问题相比别的科学的研究者较不重要;一旦克服了语言障碍(通常通过运用英语),一个法国数学家在谈到他的概念和方法时对一个中国数学家要比对一个他本国的工程师亲近得多.

数学家之间基于国籍的自然联系要比基于学派的联系少得多.虽然许多学派集中在一个国家中,但有一些国家活跃着若干学派,而且学派的影响远超出国界.当然,它们并不是在时间和空间上有明确界定的实体,但由一种持续的传统联结起来,并共有一些大师以及某些受到偏爱的主题和方法.

1800 年之前,数学家人数很少,分布不广,也没有严格意义下的门生.尽管如此,从 17 世纪中叶起,数学家之间通信频繁,数学研究几乎一直没有间断地在前进.从年代上说,最早的数学学派是法国大革命后在巴黎形成的,这是由于综合工科学校的创建提供了数学家的摇篮,直到大约 1880 年后,这根权杖才由高等师范学校接过去.在德国,高斯一直卓尔不群,但高斯之后的几代数学家在一些大学中创建了数学研究中心,其中最重要的在柏林和格丁根.在英国,大学里的数学研究在 1780 年后进入休眠状态,直到大约 1830 年由于剑桥学派的诞生才得以苏醒,这个学派 19 世纪在逻辑和代数方面特别多产,而在数学物[25]理方面一直光彩夺目直至今日.意大利在 1700 至 1850 年间只有几个孤立的数学家,但其后有一些活跃的学派成长起来,尤其是在代数几何、微分几何和泛函分析领域.

在 20 世纪,由于战争和革命,数学学派经历了许多盛衰兴败.首先是有些国家,基于极不相同的社会和政治条件,出现了难于解释的奇迹.1914 至 1918 年战争后,在苏联和波兰突然出现了第一流数学家群星汇聚的局面,而这两个国家以前只产生

过少数有国际声望的科学家.同样的奇迹出现于1939至1945年战争后的日本,尽管那里蹩脚僵死的大学系统使得日本学派中某些最好的人物流失到美国.

妨碍数学进步的有些因素是易于理解的.在法国,年青科学家群体在1914至1918年战争的中流尽了鲜血,于是法兰西学派变得围绕在它的一些老数学家周围昏昏入睡.另一方面,德国在保存它的科学家的生命和保留它的大学的伟大传统方面做得比较好,这就保证了它的数学学派具有非同寻常的影响.许多国家的大学生到德国去接受训练,尤其是年青的法国人,他们热切地努力恢复同由于庞加莱1912年去世而被遗忘的自己国家传统的联系.但是,1933年后,德国和意大利学派的繁荣被法西斯残暴地扼杀,只是到1950年以后,它们才得以重整旗鼓,而这一次是在(很有意思,情形正好倒过来)具有"布尔巴基"倾向的法国数学家的影响之下.波兰的数学学派在肉体上被消灭了,因为半数波兰数学家惨遭纳粹杀戮.直到1970年后他们才恢复原有的地位.至于英国和俄国的学派,他们尚能通过战争的艰苦磨难,没有遭到巨大损失.

最后,仔细研究一下美国各个数学学派的建立是值得的,因为它说明了即使在一个人口众多、资源丰富的国家,培植原本并不存在的研究传统是何等困难.一直到1870年,那里并未出现任何有声望的创造性数学家;大陆的开发使得抽象思维的空间所剩无几.1880年后,开始了创建数学中心的最初努力,就是邀请欧洲(尤其是英国)学者到一些大学访问和任教,创办数学期刊,以及派遣有才能的年青大学生赴欧接受欧洲数学的基础训 [26] 练.从大约1900年起,这些努力获得成功.第一个赢得国际注意的数学学派是芝加哥学派,接着在1915至1930年之间有哈佛学派和普林斯顿学派.大量欧洲数学家因极权主义政权驱逐而移居美国使美国数学力量得到意料之外的增强.正是这批人为现今土生土长的辉煌的美国学派的繁荣作出了巨大贡献,这些

学派以其在群论、代数拓扑和微分拓扑以及其他分支中的惊人发现而处于这些领域的前沿.

在印度和拉丁美洲,数学学派也在缓慢地发展,但由于政治动乱和经济动荡而遭到更多的困难.中国在经历了60年的动乱之后,看来也已进入这样的进程,从那些以前能在国外从事数学研究的人的成就判断,很快就会有某些极有才华的数学家出现在舞台上.不过,许多处于发展进程中的国家未能成功地建立持久的数学学派.此外,数学也像所有科学一样,研究的队伍必须大到一个"临界数量",除非例外情形(1900年后的斯堪的纳维亚国家和1900至1940年之间的匈牙利),国家太小就不能指望有真正的民族学派,而多多少少需要附属于人口较多的邻国的学派.

易于理解巨大的、活跃的数学学派的强大吸引力.一个只靠自己的年青数学家,面对大量的参考文献,会困惑得不知所措,很快就会灰心丧气.在一个主要的中心里,尚未入门的研究者,听取大师和资历较深的同辈以及国外蜂拥而来的访问学者的看法,很快就能进入阵地,区别哪些想法和结果对于形成他的工作的基础具有本质的意义,而哪些则只占次要地位.他能得到指导,直奔关键论著;他能获知当今的重大问题以及他们用来作为进攻武器的各种方法;他会得到警告,避开不会有什么结果的领域;他还能从他本人的研究与他同事的研究之间的始料未及的联系中受到启发和鼓舞.

由于有了这些专门的研究中心以及把整个地球联系起来的通讯网络,今天就很难再会出现像过去某些革新家所经受的那种缺乏相互了解的局面.事实上,现在只要宣布了一个重要结果,紧接着几个月中其证明就几乎会到处传播,并有很多人忙于[27] 对它进行仔细的考察和研究.

第 二 章
数学问题的性质

1．"纯粹"数学和"应用"数学

眼下"基础科学"和"应用科学"是流行名词．看来大家都接受，一门科学或一门科学的某一部分，如果其目的在于理解现象，那么它就是"基础的"．而如果它是着眼于由人类为其自身目的（好的或坏的）来控制这些现象，那么它就是"应用的"．如果说有些人由于"应用"科学所能带来的祸害而对它心怀恐惧，那么很少有人对"基础"科学的价值提出质疑，可能的话把它同由它作出的应用分离开来．

通常与"应用数学"相对，我们不说"基础数学"而说"纯粹数学"，这一事实意味着数学的这两个部分之间比起别的学科来有更多的区别．确实，如果我们稍微进一步分析一下这两个名词的含意，就会得出这样的结论：两者实际上是极不相同的思维过程，或许称为"数学"和"数学的应用"会更好些．事实上，当我们比较，例如分子生物学与临床病理学时，我们发现两者都涉及同一对象——细胞，前者考虑其内部结构，而后者考虑的是由细胞组成的人体器官的整体功能．与此相反，数学家讨论的"对象"与工程师和物理学家处理的对象根本没有相同的性质；下一章在很大程度上就是致力于弄清什么是数学"对象"．然而，我们现在就可以说，古希腊人已经清楚地理解到的是，我们的感官决不能 [29] 把握一个数或一张平面，这与看见一堆苹果或一堵墙不同．

那么数学的"应用"是由什么组成的呢?在我看来它能描述如下.它的任务在于基于物质世界某些对象的性态所遵从的一般规律,预言这些对象在给定条件下的行为.通过使物质对象对应到认定能"表示"这些物质对象的数学对象以及使控制前者的规律对应到数学对象之间的数学关系,就能构造所研究的情形的数学模型;这样原来的问题翻译为数学问题.如果能以精确或近似方式解此数学问题,就可再把所得的解翻译回去,从而"解"出原先提出的问题.

来自当今世界的一个例子是人造卫星和宇宙飞船的制导.这样的对象可表示为空间中的一个点,对该点还赋予表示其质量的一个系数即一个给定的数,点的位置可通过它关于固定坐标轴系的 3 个坐标(又有 3 个数)确定;最后通过时钟标记观测该点所经历的时间,又得到另一个数.宇宙飞船受到地球、月球、太阳以及其他行星的万有引力的作用;在相应的数学模型中,这些力表示为向量,其分量看作飞船坐标的函数.于是按照动力学定律,飞船的运动过程翻译为一个数学模型即一个微分方程组的解.数学家们应用能给出这种方程组的近似解的方法,就可得知每个时刻飞船的坐标,只带有少量误差.但在使用电子计算机以前,为得到这些数据需要进行几个月甚至几年"手工"计算.现在我们有高效能的计算机,得以几乎在瞬间即可完成这样的演算,因而我们能预先定出,为使飞船按所设计的轨道运行,应沿什么方向并以多快速度发射.

这个例子是特别简单的情形,因为它只涉及到 3 个称为飞船参数的数作为第四个参数即时间的函数的计算.通常"参数"的数目多得多,而且参数之间的关系远比牛顿定律复杂.

在 17 和 18 世纪,为力学定律和行星运行(称为"天体力学")提供数学模型取得了巨大成就,它与观测结果相符到令人[30]赞叹的地步.事实上,这些模型正是数学最早的实实在在富有成效的应用.它所用的定理属于微积分的所谓"初等"部分,它唤起

了同时代人的钦佩,但现在它已丧失新奇的魅力,中学最后一年或大学前两年都讲授这些内容.于是数学家们觉得,如哈代所说,它"单调枯燥,令人厌烦,缺乏美学价值".但是,如我们会在第四章中看到的,对于一个数学问题来说,即使当它解决之后,也不引起大量其他问题,这是十分罕见的.就动力学微分方程组而言,在一段时期的停滞之后,从 1880 年起,随着 H·庞加莱的工作,出现了一个崭新的数学领域,称为"动力系统"理论,它富含困难深刻的问题,当前很多数学家比以往任何时候都更热切地致力于这方面的研究.

2. 理论物理学与数学

数学对物理学的应用在 19 世纪得到了扩展,在 20 世纪它仍在进一步扩展,但其特征有了变化.在为流体动力学、弹性论、电磁学、热力学以及稍后为相对论和量子力学等新理论构造数学模型时,数学家发现他们面对的是远比天体力学数学模型更加难以对付的问题,尤其是它们涉及所谓偏微分方程.最简单的例子之一是"弦振动"方程

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0,$$

其中 x 是弦上点的位置,t 是时间,$u(x,t)$是弦上位置为 x 的点在时刻 t 离开该位置的距离.

尽管某些这类模型可追溯到 18 世纪,但那个时代的数学家在解他们提出的这些问题时多少无能为力.为获得可为物理学家应用的重要结果,耗费了 19 世纪分析学家的全部努力.然而至今这些问题中的许多问题仍然只是部分得解,这就不断产生 [31] 新的研究和新的解决方法.

因此我们能毫不犹豫地说,正是物理学的需要促使数学家创建数学的一个新分支,它称为泛函分析.现在它是数学中很可观的一个领域,并在不断前进并衍生出各种新的分支.值得注意

的是，有时把专属于物理学的某些概念(能量，“最小作用量”原理，等等)移植为数学模型，能得到解泛函分析中某些方程的一般方法.

还有，最近把概率论应用于统计学和物理学导致这一理论中的许多工作；同样，运筹学和自动机理论促进了代数学某些分支的研究.

3. 经典时代数学的应用

牛顿力学及其在天文学中的巨大成功，尤其是克莱罗预言哈雷彗星的返回(精确到一个月以内)，在欧洲知识界中激起强烈的印象，甚至包括像伏尔泰这样对数学绝对一无所知的文人①.至于对18世纪的数学家如克莱罗、达朗贝尔和拉普拉斯，数学的这些非凡应用以及随后的其他应用，导致他们设想数学研究的基本目标是为力学和物理学提供模型；任何不满足这些条件的数学分支则被认为是无益的、可以忽略的.当欧拉写信给克莱罗告知他在数论中得到的一些结果并说对此他一直干了14年时，克莱罗只作了这样的评论：“那必定是一个充满荆棘的题材”，跟着立即改变话题，转向他感兴趣的计算行星扰动的主题([6],PP.129).

17世纪之前，数学的应用远没有像牛顿力学所引起的惊叹那种近于神奇的特性.事实上，当时数学家能为之提供模型并对[32]其现象作出合理解释的，只有4种学科.它们都归功于古希腊数学家，而此后直到16世纪前没有出现过新的思想.这4门科学是：镜面光学(或“反射光学”)，静力学，浮体平衡学，这三者的肇始似乎都不早于公元前4世纪，还有天文学，有文字记载的第一步是在公元前6世纪.

---

① 他曾写道，他永远也不能理解，为什么一个角的正弦不与该角度成正比.

这些应用都仅仅基于几何学,从而被看作几何学的附属学科①.由于缺乏能使他们描述任意运动的数学工具,诸如只有凭借微积分才能合用的数学工具,古希腊人为"追踪"天球中行星的运动,就从一个先验的观念开始,即对于天体,只有绕一个轴(如果在一平面上,则是绕一个点)的均匀旋转才是可接受的.他们把它归之于至高无上的"完美性",无疑这是关于这些天体"神圣性"的神学观念的残余.众所周知,这导致他们想象出相互运动的球面或"本轮"(其中心画出别的圆的圆)体系,为使之符合以后变得越来越精确的观测,这个体系也变得越来越复杂.

但在进行这种最后注定要失败的尝试之前,古希腊人对地球以及其直径可被肉眼清楚地估计到的两个天体即月球和太阳的形状和大小,形成了虽然近似但却基本上正确的观念.为此他们只运用他们的几何学,以独创的方式把它用之于观测.它虽然简单,却是以罕见的机巧来选取的.我要在这里概述这些第一流成就,在精神上它同别的民族对待同一现象的态度极不相同.甚至巴比伦人,他们的天文学是最发达的,但除了注意到这些天体的运行(它们的升起和降落,它们的会合和相蚀,等等)以及发现影响它们的多重周期性之外,没有再前进一步;我们没有发现任何迹象,说明在他们的思考中有过为解释这些现象而设计的几何模型.

我们必须记得,在古代世界和中世纪,可供天文观测用的测量仪器只有长度可变的直尺,它称为"对准盘",绕着一个有刻度的圆的圆心作回转运动,使得能通过置于该装置远端的一些洞进行目测,确定通向天球而上两个点的两条直线之间的夹角,其精确程度视情况而异,用最好的这类仪器或许可达到圆弧上一 [33] 个分那样的精度.于是,每个模型都得适应这种角度的测量.

① 在1900年前后重新发现的一篇简短论著中,阿基米德解释过,为了想出他关于面积、体积和重心的几何定理,他如何把所讨论的图形分成"薄片"并按其重量来比较这些薄片.

这方面的第一步是由公元前 6 世纪的毕达哥拉斯学派跨出的，他们确信地球、月球和太阳是球形体，而月球的形态是太阳照亮部分变动的结果，这种变动分别起因于太阳和月球绕他们设想为固定的地球的转动. 他们知道月蚀起因于月球通过地球的阴影锥(图 1)；我们看见这个锥在月球上的痕迹是一圆弧(图 2)这个事实只能用地球表面是球面的假设来说明.

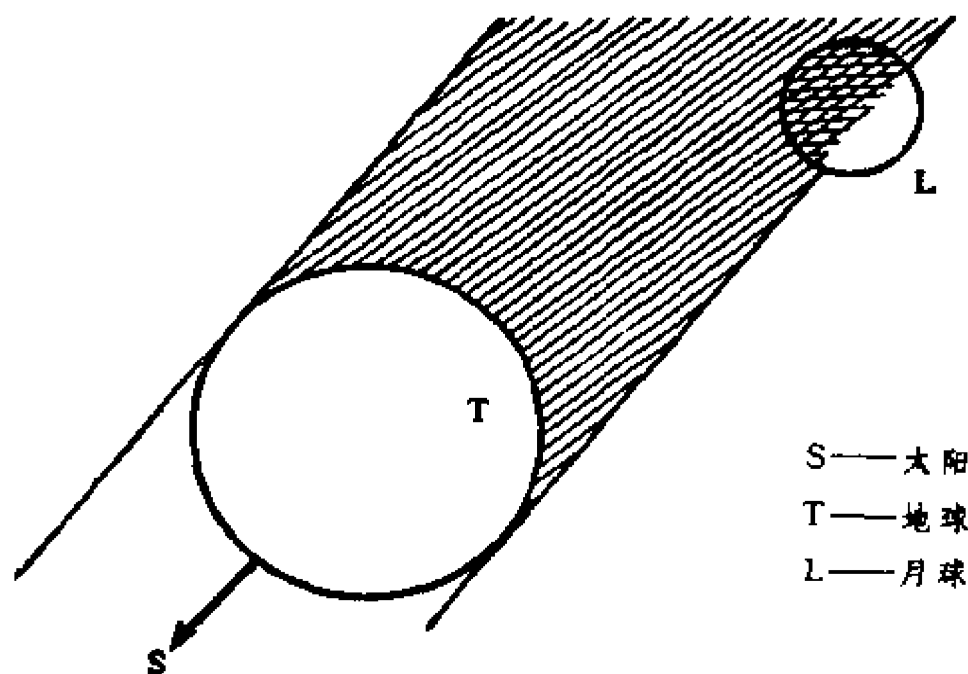


图 1

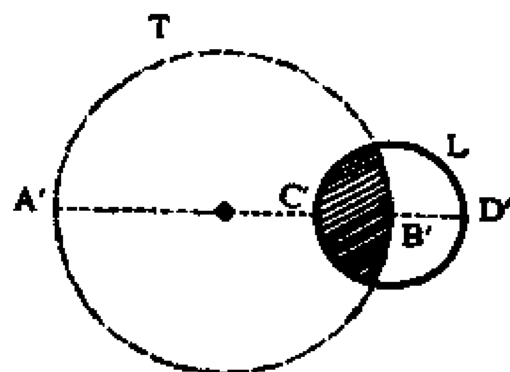


图 2

对月球进行的另一项观测使他们发现距离 TS 远远大于距离 TL①.事实上,在第一象限内,把月球的明亮部分与阴影部分分开的平面通过 T,因而直线 TL 垂直于 LS(图 3);由此

$$\frac{TL}{TS}=\cos\alpha .$$

角 α 非常接近于90°(与90°之差小于1分),因而比 TL/TS 的阶为0.003.由于瞄准太阳很困难,所以古希腊人设想 α 近似于87°,因此对于 TS 所得的值过小.

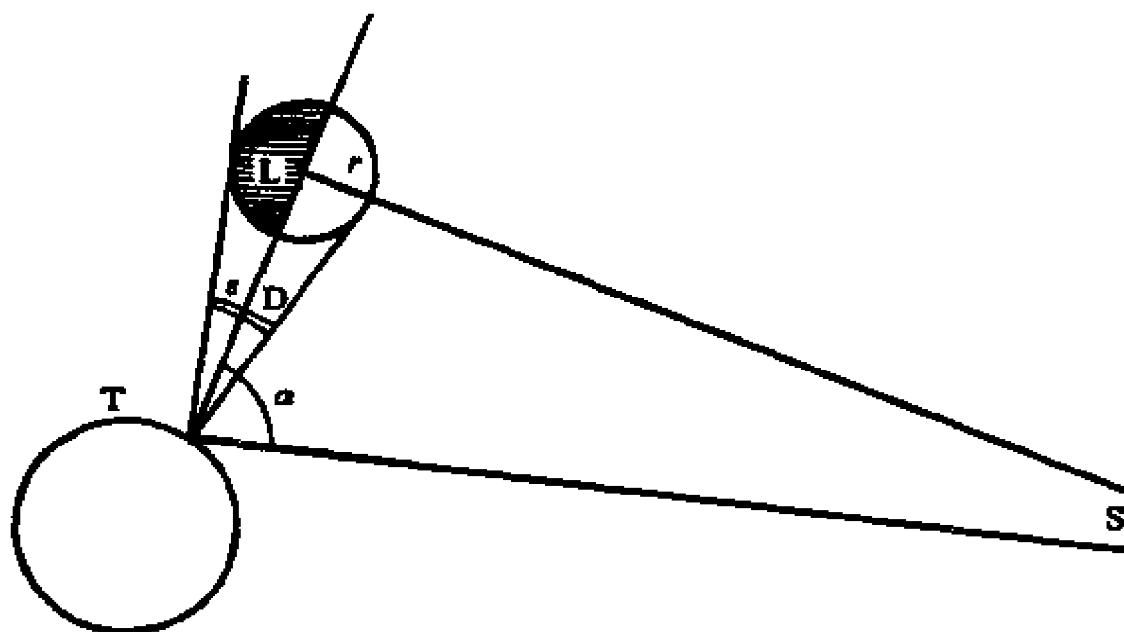


图 3

这样,如果我们把地球表面上所有点处收到的从太阳中心发出的光线看作平行线,就只会产生极小的误差.这使得爱拉托斯特尼于公元前大约300年得出地球半径的第一个近似值,其

① 在图1到图3中,T,S,L 是取所说三个天体中心或取在其表面上的点.这些点取法的变动产生的误差小于因古希腊人所用仪器不够完善而引起的测量角度的误差.

推理过程如下.在埃及,亚历山大与赛伊尼(现阿斯旺)多少位于相同的子午线上(图 4).在夏至这天,中午时太阳正好在赛伊尼头顶上,同时太阳光线与亚历山大处的铅垂线作成角 β.由于角 β 是地球中心 O 连接亚历山大和赛伊尼的两半径构成的角,所以两地之间子午线段为 $d=R\beta$(β 取为弧度),由此 $R=d/\beta$,而爱拉托斯特尼关于距离 d 有一个粗略估值.

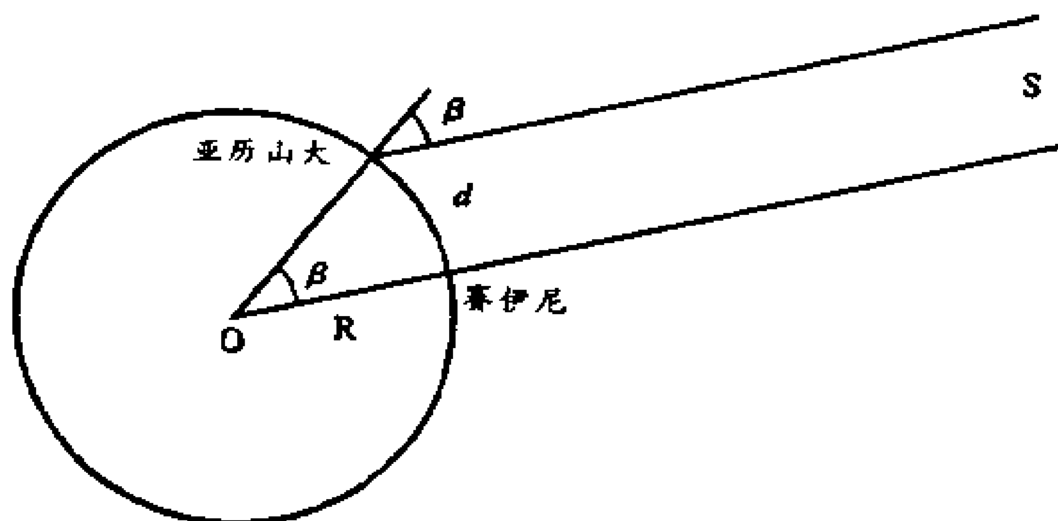


图 4

有了 R 的值,如果我们知道地球半径与月球半径之比 $A'B'/C'D'$(图 2),就能推出地球到月球的距离 TL.通过观测地球在月球上阴影的曲率,古希腊人判断这个比值近似为 3;事实上此比值为 3.7.如果我们测出月盘所张的角 ε,那么就有 [34] (图 3)$D=2r/\varepsilon$,其中 r 是月球的半径.

应当注意,我们方才描述的数学应用,犹如牛顿及其 18 世纪继承者的天体力学,涉及的是"基础"科学.在人造地球卫星发明之前,它们本身并无直接的"用途".尽管如此,这门科学所获得的普遍赞美,是古希腊人以来绵延不断的传统的一个例证,这种传统,如亚里士多德所说(《后分析篇》,981ᵇ),就是不把科学

的价值归之于实用或获利的目的[1].在这里我们接触到人类特征之一——好奇心,它在儿童身上就已非常显见,尽管许多成人为日常工作和智力惰性所淹没,看来已经丧失好奇心.

当然,古希腊人从来没有想到把数或几何图形的性质从它们在天文学中的应用里分离出来.相反,从公元前6世纪的毕达哥拉斯学派起,整数常带有神秘主义的色彩,在柏拉图那里这一点仍很显见,它一直延续到今天"命理学"[2]的异想天开之中.

4. 功利主义的责难

正是在围绕达朗贝尔的圈子(其中有布丰,尤其有狄德罗)中,关于数学真正地位的争论走得很远,达到了这样一点:数学家应当瞄准的唯一目标,是数学对于其他科学的应用.对狄德罗来说,数学有过它的鼎盛时代,现在它对经验不能增添任何东西,它不是"使哲学普及"而只能"在自然和人类之间挂起一幔幕布".我们看到,"文化革命"的原理并非近来才有的! 不像伏尔泰,狄德罗涉猎过数学,但他应当知道,他从未在数学中做过独创性的工作.我们是否应当把他的责难归之于这一点? 正如丰特奈尔在1699年就已指出的:"我们通常总把我们不理解的事物称作无用.这是一种报复心理,而因为数学和物理学一般不为人所理解,因此它们就被宣布为无用的."

为对狄德罗以及那些时至今日仍在重复他对数学的抨击的人公平起见,必须承认,直到20世纪初,力学和理论物理学对于"有用的"技术发明几乎毫无影响:"简单机械"(杠杆、滑轮、螺旋

---

① 无论从所引亚里士多德全集希腊文版的编号 981^b 或所引内容看,这段话似出自《形而上学》;参见亚里士多德,形而上学,商务印书馆,1991,PP.3.——译注

② 命理学(numerology)根据毕达哥拉斯万物都可归为数的思想,用数字解释人的性格或占卜祸福.——译注

[36] 等等)可追溯到遥远的古代;蒸汽机先于热力学定律,甚至内燃机和航空也来自“修修弄弄”而不是来自理论.只是到了今天,离开“基础”科学(因而包括数学)的结果,技术才一事无成.这样,设想把数学家的工作限制在狄德罗意义下可以“普及”的工作上的革命家,不得不迅速开倒车.

对于一些小册子中不时迸发的攻击“纯粹”数学的恼怒,仍然是很难理解的.有许许多多别的学科,例如宇宙论,史前学,考古学,很清楚它们是毫无“用处”的,但却没有人责难它们.相反,当政府缩减对于它们的拨款时,很容易激起支持这些学科的舆论.我们是否又应援引丰特奈尔的名言?

5. 时髦的说教

轻视“无用”数学的人应当看到,在当代数学的大量成果中,相当大的一部分在今天明显地没有可以想象到的“应用”.他们通过坚定不移地断定即使是这些部分也会在未来某一天变得“有用”来避开这个困难.由于这样的论断涉及不能确定的将来,它只不过是一种说教,既不能驳倒,也不能证实.这方面无法辩驳的论据是,开普勒在圆锥曲线理论中发现行星绕太阳运行的数学模型,而这一理论是阿波罗尼奥斯在1800年之前论述的,当时并无把它“应用”于任何一件事物的一丁点儿企图.最近又能添加两个类似的情形:在60年前构想的黎曼几何中发现了广义相对论的模型以及原子和“基本粒子”结构,其中群论和希尔伯特空间起着重要作用.

然而这里有数论中3条最著名的定理:

Ⅰ (欧几里得,约公元前300年) 存在无穷多个素数(或按其原来表述,每个自然数小于某个素数).

Ⅱ (拉格朗日,1770年) 每个自然数能写成一个、两个、三个或四个自然数平方之和的形式(例如:$7 = 2^2 + 1^2 + 1^2 + 1^2$).

[37] Ⅲ (维诺格拉多夫,1937年) 存在(很大的)自然数 N,使

得每个大于 N 的奇数是一素数或是 3 个奇素数之和①.

这些结论陈述的都是涉及自然数的无穷性的性质.如果我们减弱这些陈述,譬如说,限于小于 10^{1000}的自然数,我们仍会得到相应的定理,但绝大多数数学家就会对它们不感兴趣.很难设想一种物理学理论会需要考虑无穷数集.

还有另外一些由某些科学史家所表述而使数学家感到惊讶的观点.这些人并不满足于这样的数学史论著:它们阐述过去的各种数学思想,力求理解它们之间的内在联系和相互影响.按照这些人的看法,还有必要来"解释"数学家为什么选择这个或那个研究方向以及他们是怎样得出他们的结论的.我得承认我不理解这种看法能有什么意义.一个创造性头脑中发生的过程决不会有合理的"解释",这一点在数学中比任何其他心智活动中尤为突出.我们仅仅知道它需要有一段不出成果的研究阶段,有时这段时间会相当长,接着突然"柳暗花明",显露出来的东西得到"定形".所有这一切,都由庞加莱以其切身体验很好地描述过([9]).

至于数学家要着手解决的问题的起源,则几乎总是在同别的数学家的接触中寻求.这种接触或是通过研读经典论著或教学性图书,或是通过谈话或通信,或是通过听取讲演.但这对于那些科学史家还不够,他们是要对数学家在其自身学科中形成的概念加以评判,并竭力进行他们的普遍"解释"即社会环境的

① 这类分解的一个例子是 $33 = 23 + 7 + 3 = 11 + 11 + 11$.人们猜想这种分解对任意的奇数都成立,于是我们应当对小于 N 的整数进行验证;但 N 实在太大,即使最有威力的计算机也不可能完成这种验证工作.如果有读者设想这样的论断在数学中很少见,那么我就要请他查阅 L·E·迪克森所著的篇幅达 1289 页的《数论史》(History of the theory of numbers, 2 vols., Carnegie Institute, Washington, D.C., 1920).书中详尽无遗地叙述了从肇始时期至 1918 年所有与数论有关的文章.如果要论述嗣后在这一数学领域中所做的全部工作,那就得写另一本同样篇幅的书.

影响．这是知识界许多人宣布的一个熟知的教条，希望以此来贬[38]低个别人物的贡献和创造力并同他们所谓的“尖子主义”进行战斗．我没有能力来判断这个教条在其他科学中的正确性，尽管我不能看出牛顿或爱因斯坦所生活的社会如何影响他们的发现．但在数学中，不算为其他科学提供模型的那些部分，在我看来，鉴于过去时代数学家的工作方式和当代数学家的表现，这个教条是十足谬误的．怎么能设想像托勒密王朝时代的亚历山大，路易十五时代的巴黎或腓特烈二世时代的柏林，还有斯大林时代的苏联这样不同的社会环境对上面引述的欧几里得、拉格朗日还有维诺格拉多夫的数论定理会有什么共同的东西？

这个教条还有另一方面，即宣称它看出了科学研究人员期望自己的工作对社会“有用”的动机．对此我只引用哈代所说的话([7],PP. 19)：“如果一位数学家、一位化学家或甚至一位生理学家告诉我，他的工作动力是使人类得益的意愿，那我是不会相信他的(如果我相信他，我也不会因此对他有更高的评价)．”我确信，绝大多数数学家都这么想，尽管他们可能由于害怕被占主导地位的“知识界”耻笑而不公开这样说．

6. 小结

让我们把本章长长的讨论用几句话总结一下．

Ⅰ 数学中有一个完整的重要领域，它产生于为其他科学提供模型，毫无疑问不能贬低这个领域．尽管如此，它在当代数学中所占的比重肯定不超过 30% 到 40%；如果我们浏览《数学评论》月刊(它对数学及其最重要应用中发表的论著给出概略分析)，就能容易地看出这一点．

Ⅱ 我们能同意哈代所说([7],PP. 20)，激励数学家做研究的主要动力是智力上的好奇心，是谜团的吸引力，是穷究真理的需要；正如希尔伯特所说：“问题就在那里，你必须解决它”．我们会在第四章中看到，这种永不满足的激情如何导致大量问题离

奇地绽开嫩芽，使得数学家们永不止息地投身于这些问题的研究.

当然，在这之后还有其他动机，它们是毫不丢脸的，是数学家与许多其他人群分享的，诸如在身后留下某些恒久作品的愿 [39] 望，肯定也有这样的抱负，即在生前就获得对他们所做工作的尊重.

Ⅲ 数学家清楚地意识到，这样的尊重只能来自他的同行. 他生活在一个封闭的圈子里，几乎与外部世界不通音讯，而且除为其他科学提供模型的那些部分外，他正是在这个圈子里得到自己研究的问题.

Ⅳ 不管一个数学家对自己的工作可能有多高的看法，只有他的同行的评定才会建立其价值. 显然他的结果必须是“不平凡的”. 他们克服的困难愈大，他们所解决的问题此前被不成功地尝试的次数愈多，他们获得的赞美就愈热烈. 除此之外，没有适用于一切时代和所有数学家的评价标准. “一般性”或“深度”这样的词是常常用到的，但其意义对于使用这些词的人并不总是相同的. 数学中也有一时流行的风尚，它把某些领域暂时拔高而牺牲另外一些领域. 这种品味上的歧异使人想起围绕艺术作品的争论，事实上数学家之间也常讨论一条定理是否为更“美的”或不那么“美的”. 这一点总是引起从事其他学科的人们的惊奇：对于他们，唯一的标准是一门理论或一个公式的“真实性”，即它在多大程度上能满意地说明观测到的现象. 在数学中，在根据普遍接受的逻辑规则已经证明的意义下(见第三、六两章)，所有结论都是“真的”；数学中没有未加证明的论断的位置. 因此，评价数学工作就需要另外的标准，这些标准不可避免地具有主观色彩，这就使得有些人说数学更多地是一门艺术而不是一门科学. [40]

第 三 章
经典数学的对象和方法

尽管所有古代文明为满足日常生活需要，不得不发展算术演算和空间测量手段，但只有古希腊人从公元前 6 世纪开始，想到分析这些手段背后的一连串推理，从而创造了一种崭新的思维方式. 在本章中，我们试图弄清古希腊数学的本质特征以及文艺复兴时代至 18 世纪末的数学家所带来的出乎意料的成果如此辉煌的发展.

在 § 2 和 § 3 中，我们将集中讨论古希腊数学的两个基本特征：

1）从未经证明的命题、公理和公设出发通过相继的逻辑推断作出证明的观念. 必须强调，只有借助熟练运用逻辑的技巧，这个观念才能付诸实践，而这种技巧是由古希腊哲学学派培育的. 一个特别引人注目的例子是"反证法"原理，它是由逻辑学家提炼的工具，却成为数学推理的支柱之一.

2）数学家关注的对象同实际演算中所用的东西具有相同的名称：数、几何图形、量. 但从柏拉图时代起，数学家已意识到，在这些名称下，他们是在对完全不同的实体即非物质的实体进行推理，这种非物质实体是从人们的感官所能感知的对象"通过抽象"得到的，而后者只是前者的"形象".

§ 4 讨论几何图形的概念，我们在该节中要揭示，通过公理赋予几何学"抽象"对象的性质如何使得这些对象同它们的"形[41]象"之间产生深刻的差别，以及由于这种差别在寻求定义这些对

象的适当词汇方面出现了什么样的困难.为了能弄清这些观念,我们将不在这里停下来去详细追述这些概念的历史变迁(在第六章中我们会回到这个论题),而是一下跳到帕施和希尔伯特于19世纪末所跨的一步.他们两人吸取了欧几里得公理方法的独创精神,填补了他的空白,通过断言正是制约数学对象的公理定义了这些对象,一劳永逸地避开了这些困难.

在§5中,我们以同样的方式致力于讨论其"形象"是人们感官所感知的实在的数和量的数学对象.整数概念的"抽象"特征总是在古希腊算术中显示出来,而欧几里得关于整数的可除性和素数的阐述仍然具有重大价值,至今还在讲授.虽然不像欧几里得的几何学,并未把这些内容置于公理式理论的形式中,但为使它获得此种形式只需添加少许东西(第六章).

与此相反,不可公度量的发现使古希腊人的量的度量概念出现了危机.事实上,看来从前毕达哥拉斯学派总是认为,一旦对某类量选定了一个单位,那么每个这类量关于此单位必是"可公度的"——我们喜欢说成它的度量是一有理数.为克服这个困难,古希腊人创造了新的数学对象,即同类量之间的比.这些对象是以这样的方式公理式地定义的:一旦对一类量选定一个单位,那么同一类中的量对此单位之比构成我们所说的正实数集的一个部分.这个部分含有有理数,还含有某些无理数,但所讨论的主题不允许我们确切地划定其界限.

无疑是由于哲学理由,柏拉图学派的数学家遵守关于运用三类几何量——长度、面积、体积的禁忌.例如,不能把长度的度量与面积的度量相加,两个长度度量之积是一个面积度量(三个长度度量之积是一个体积度量)而不是长度度量.虽然几何学可以容纳这些限制,但当我们对实数进行演算时,这些限制使代数运算变得不可能.为使采纳这类运算得到认可,曾经需要笛卡儿的权威,尽管某些数学家在前几个世纪中早已提出过这些运算.嗣后同类量之间的"比"就等同于实数,不必再特别指明所考虑

[42] 的类.

在§7和§8中,我们要揭示,连同中世纪和文艺复兴时代发明的方便记号,上述改革不仅使代数的发展成为可能,而且也使坐标方法这一极为重要的发现成为可能.这样,一方面,它为欧几里得几何学提供了代数模型,另一方面,就有可能出现实变量的实值函数这个一般概念,这是古希腊人始终未曾想到过的概念.

最后,在§6和§9中引进两个基本数学概念——逼近以及由之导出的极限.古希腊数学家通常通过几何"作图"来解代数问题,例如,欧几里得通过圆与直线之交作出一个"比"的平方根,而梅内克缪斯通过两条圆锥曲线之交来作立方根.我们在欧几里得那里还找到另一种定义平面上非多边形图形面积的度量的想法:把所给图形置于两列多边形之间,而这两列多边形面积之差趋于零.阿基米德反复使用这一想法;这种想法在17世纪得到推广,从而有可能证明,例如,n 次($n\geqslant 4$)根的存在性,这是古希腊人用作图手段达不到的.当然,为验证这些方法,有必要引进一个公理,直到19世纪柯西以"闭区间套公理"的名称使它最终明显出现之前,这个公理一直是没有得到清楚指明的.这一公理连同欧几里得的那些公理,完成了所有实数构成的集合的公理式定义.它为微积分提供了坚实的基础;微积分是于17世纪发明的学科,它注定要成为纯粹数学及其应用的最强有力的工具.

1. 准数学观念的诞生

在当代社会,孩提时代就已获得数和度量的观念.从十二三岁起它们就显得如此"自然",以致其使用成为无意识的事情.然而,皮亚杰通过实验揭示,虽然"任意"一个自然数的观念可能在很年幼时即能掌握,但某些类型的量诸如体积或重量就是另一回事,即使12岁的孩子在比较两个属于相同的这些类型的量时仍然会有概念上的困难.民族学家碰到过一些原始社会,超过若

干单位的整数便没有名称，当然就更不能用于计算. [43]

由埃及和巴比伦这些最早的东方文明流传下来的记载过于零碎，不能使我们领会这些文明中算术或初步的几何学得以形成的道路.从公元前第二个一千年起，它们已显得充分地得到发展.当然，我们不在这里讨论抽象的构思，而是讨论专门的书记阶层传下来的一些诀窍，它们是专为解决达到一定高度的农业社会结构所提出的实际问题——交换、租金、诉讼以及财产分割等问题而设计的①.

对我们试图达到的目的而言，没有必要详细阐述留传至今的文件中所解的数学问题.我们只想提一下，在算术方面，他们已有分数、算术数列或许还有几何数列以及"三分律"知识.在巴比伦人那里我们甚至能发现等价于二次方程的问题的解.例如，有一块泥板，上画一正方形，并有下述文字："若把正方形边长加其面积得 3/4，其边为何?"书记人员想的是一个方程，我们可把它写为

$$x^2+x=\frac{3}{4}.$$

对此书记员以与我们会采取的同样方法解该问题：他把方程两边同加 1/4，从而发现 $x+1/2$ 的平方为 1，由此他得到 $x=1/2$.

在平面几何领域中，可能连同一些工具，诸如陶工旋盘、测量人员的描图尺和石匠的直角器的使用，已经知道矩形、三角形、梯形、直角和圆等这样一些几何图形.巴比伦人还有相似这种想法，他们有一段文字说，有一楼梯，每一阶的高度与宽度之比与此楼梯的高度与水平投影之比相同(图 5).另一方面，古希腊人把测量金字塔高度的一种方法归功于泰勒斯；但无庸置疑，

① 古希腊作家强调这样的事实：由于尼罗河泛滥改变田地的界限，必须求助具有专门技能的书记人员，他们受命掌管公式，使每个人在泛滥后能重新获得与他原先拥有的土地相当的一片田地.

古埃及人已知道这种方法：观测金字塔影子的长度，金字塔高度
[44] 与其影子长度之比等于一根棒高度与其影子长度之比(图 6).
至于立体几何，现存的一些宏伟建筑说明建筑师和石匠有一点
经验性知识，对此我们难以说清楚.

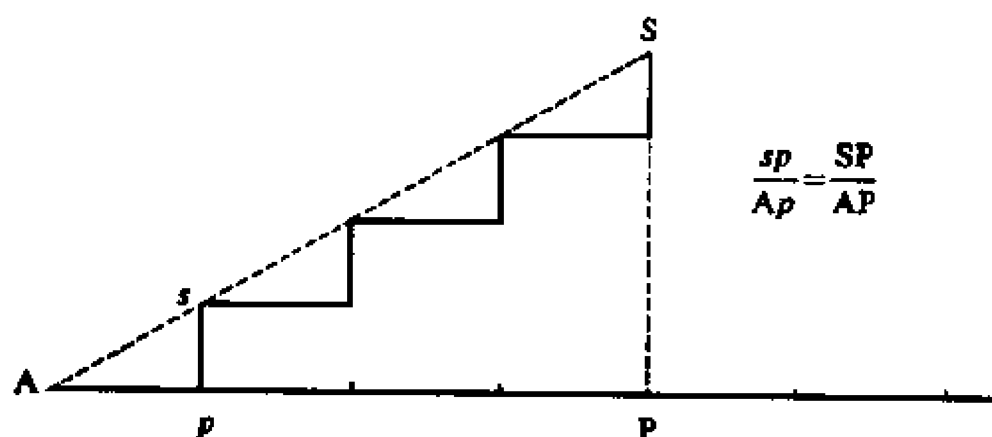


图 5

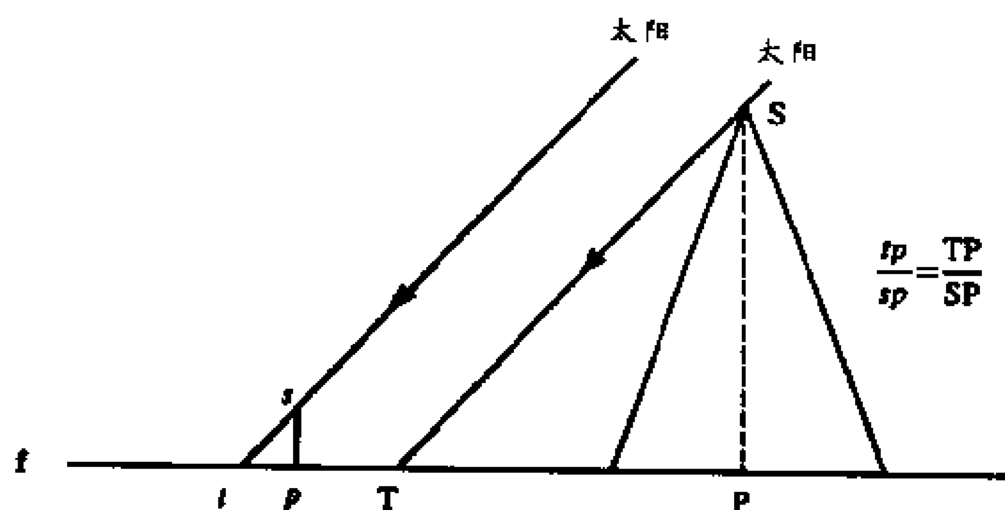


图 6

所出现的概念涉及的都只是具体的对象：对一堆物体计数，对可以相加或相减的量例如长度、面积、体积、重量、角度等加以测量. 每种这样的量都取了一个单位而且还常取其倍数和约数.

上面提到的诀窍所涉及的例子中，数据都是已指定的；它们是一些计算程序，例如计算形状和大小为已知的图形如等腰三角形、矩形、梯形或圆的面积[①]，但没有给出一般的验证. 自然，我们找不到公式——在我们所理解的这个词的意义下即对任意的或不特意指定的数据都成立的公式. 计算程序的一般性，只有当给出一系列数据变动的例子时，才能加以猜测.

2. 证明的思想

留传到我们手中的公元前 7 世纪和 6 世纪古希腊作家一些著作的片段能使我们在其中辨别出我们所说的逻辑演绎即一系列推理——后来汇编为各种三段论法[②]——的例子，它驱使一个对话者一旦接受断言 P，就得接受断言 Q[③]；但没有任何别的古代文明留下在此之前任何类似的文献. 大家知道，从公元前 5 世纪起，正如智者派的论著和柏拉图对话录的片断所表明的，古希腊思想家中出现了一些大师，他们善于把一个论证排成逻辑演绎顺序. 他们发现这些推理能以任何人类活动作为对象，尤其是以算术和几何诀窍作为对象，无疑这些诀窍来自古埃及文明和巴比伦文明. 它们成为把定理连结在一起的证明.

传说第一批定理始于公元前 7 世纪末的泰勒斯，但这只是通过后来的文本得知的，而且也不知道它们的证明，如果存在这样的证明的话. 然而，大家承认，毕达哥拉斯学派的各种定理——其中当然有熟知的“毕达哥拉斯定理”——是伴有证明的，尽管其性质仍然并不清楚. 包含证明的第一批文字只能在柏

① 对于某些多边形，它们只给出面积的近似值；另一方面，对于圆，某些古埃及文本给出 π 的较好近似值 3.16.

② 逻辑学家设计了各种三段论，其中有些从未被数学家用过.

③ 用字母来标志未特别指明的命题这种做法可追溯到亚里士多德.

拉图和亚里士多德的著作中找到.

在称为《米诺篇》的对话中,苏格拉底让一个年青的未受过教育的奴隶作一个正方形,使其面积为给定正方形 $ABCD$ 的面积的两倍(图 7).那孩子先说把边长加倍即可作出,苏格拉底告诉他这个正方形的面积不是 $ABCD$ 面积的两倍而是四倍.苏格拉底然后让他画一正方形 $A'B'C'D'$,每边等于 $ABCD$ 的对角线,并证明此正方形具有所需的性质.这是把毕达哥拉斯定理应用于等腰直角三角形这种特殊情形.其证明是先说明正方形 $ABCD$ 能被它的对角线分成四个全等的三角形,其中每个三角形,例如三角形 OAB,全等于在 OAB 的斜边上向另一侧作的三角形 $A'AB$;因此我们得到 8 个全等于 OAB 的三角形,而这 8 个 [46] 三角形合起来即为正方形 $A'B'C'D'$.为让人确信,欧几里得稍后从一系列先前的定理推出三角形 OAB 与 $A'AB$ 全等.苏格拉底在这里只满足于让他的对话者承认这一事实.

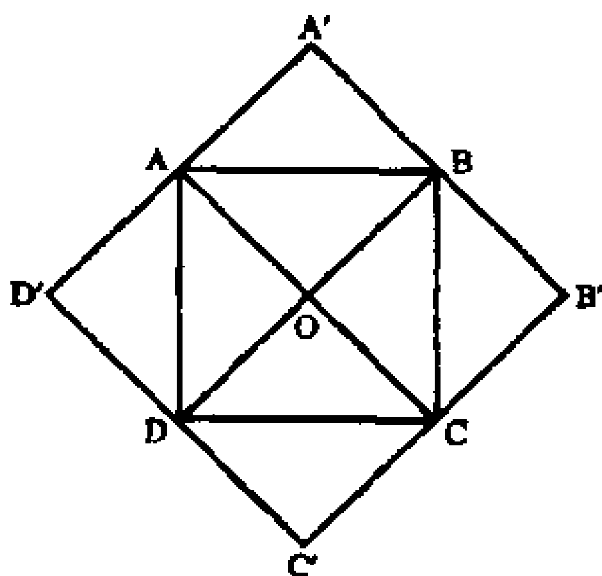


图 7

第二个证明是亚里士多德记录下来的,它来自毕达哥拉斯学派,也涉及等腰直角三角形这个图形.它是"反证法"的已知的

最早例子,而这种证法后来成为数学的一个基本工具①.它还是关于不可能性的论断的最早例子.这条定理说,等腰直角三角形的斜边与其一直角边之比不可能是分数 p/q(其中 p,q 是整数).事实上,由毕达哥拉斯定理,这样的分数会有性质 $(p/q)^2=2$.我们可以假定所述问题"简化为"p 和 q 不同时为偶数的情形:如果 p,q 两者都能被 2 整除,则可用 2 除 p,q 而不改变 p/q 的值,必要时重复做若干次,就能达到"简化"情形.现在 $p^2=2q^2$,又由于奇数的平方 $(2n+1)^2=4(n^2+n)+1$ 是奇数,所以 p 必是偶数,设 $p=2p'$.现在 $4p'^2=2q^2$,从而 $q^2=2p'^2$,因此 q 也必是偶数:我们得到了错误的结论,因此我们由之出发的假定是站不住脚的.这里进行推理的出发点是奇数和偶数的基本性质,这也是毕达哥拉斯学派所喜爱的.在欧几里得《几何原本》第Ⅸ卷中也能发现同样的偏爱,尽管那时这个主题已变得"平淡无奇". [47]

3. 公理和定义

证明必须按一系列推理来进行的方式由柏拉图在《理想国》②中用一段著名的话(该书第六卷,510,c,d)很好地描述:

"……研究几何学和算术……的人从假定奇数和偶数或各种图形以及三种角……开始.他们把这些东西看成已知的;他们把这些东西作为假设,并不觉得需要对自己或别人作任何说明,而把它们看作自明的.然后,他们从这些假设出发,通过一系列首尾一贯的步骤,最后达到所寻求的结论."

各个时代的数学家一直保留了这种方法.但是关于柏拉图所说的"假设"的性质(要注意他从未用过"真实"这个词)以及它们所应用的实体,从柏拉图时代以来,数学家和哲学家们从未停

① 有些人把这个证明的出现看作数学的真正诞生.

② 中译本:柏拉图,理想国,商务印书馆,1994.——译注

止过深入思索.这就是称为"数学基础"的问题.

柏拉图的许多对话,如其题目所示,试图阐明流行语言中许多其含意并无清晰概念的"抽象的"词,诸如美丽,勇敢,热爱,虔敬,正义,德行,等等.同样,数学家必须对出现于最初"假设"中的基本概念——图形、位置、量、数量、度量等词的涵义加以整理.我要直截了当地说,直到19世纪末以前,数学家们在这方面并未完全取得成功.到了19世纪末,这段长长的历史才告结束,我们将在本章和下章中回顾其中一些最引人注目的事件.

如果上面这些词是用来指明属于感觉经验的概念,那它们本来不会比这些经验本身更成问题.但是,在上面援引的《理想国》那段文字的结尾,柏拉图立即小心地解释(第六卷,510,d),数学家们"利用各种可见的图形,讨论它们,但他们思考的实际上是这些图形成为其形象的那些原始的对象.……他们试图看到绝对图形,那是只有通过思维才能看到的对象."

至于进一步的说明,我们只须注意《米诺篇》中这样的情景,[48] 在那里柏拉图显然暗示,要明白苏格拉底的陈述,其实与他所画的——可能画在沙上的图形无关.实际上他可能在"实验地"证明三角形 OAB 与 $A'AB$ 全等上出现了困难.

柏拉图把他的描述和考虑同他关于理念的理论联结起来,这一点与我们要讨论的内容无关.亚里士多德尽管不同意柏拉图的理念理论,但仍进一步肯定《理想国》中前引的段落.他写道(《形而上学》①,k3,$1062^{a}20-{}^{b}3$):

"数学家研究通过抽象得到的事物,因为他们先消除一切可感觉的素质,如轻重、冷暖、软硬等等,只剩下量性和连续性,后者可能沿一个、两个或三个方向想象②."

第一批"数学对象"就是以这种方式引入的.但为理解它们

① 中译本:亚里士多德,形而上学,商务印书馆,1991.

② 此处显然暗指空间的3个维数.

是什么以及关于它们说什么才合理，却经历了无数困难，直到19世纪末为止.

虽然我们有相当大一部分柏拉图和亚里士多德的论著，但欧几里得之前的数学著作，却几乎没有什么留传下来. 因此我们只能从欧几里得的《几何原本》获得关于公元前5世纪和4世纪古希腊数学家的观念的比较准确的信息，尽管这本书看来是把不同时代的几本著作汇编而成.

作为柏拉图描述的过程的出发点的"假设"，其大部分又是从较早的"假设"导出的. 例如，苏格拉底在《米诺篇》中的推理，是欧几里得书中一条定理(第Ⅰ卷，命题34)的特殊情形，该定理证明平行四边形 $ABCD$ 中两个三角形 ABD，BCD 全等(图8). 为此他应用"三角形全等的一种情形"(第Ⅰ卷，命题26)，注意到这两个三角形有一公共边 BD，且由第Ⅰ卷命题29，我们还有角的相等关系$\angle ABD=\angle BDC$，$\angle ADB=\angle DBC$.

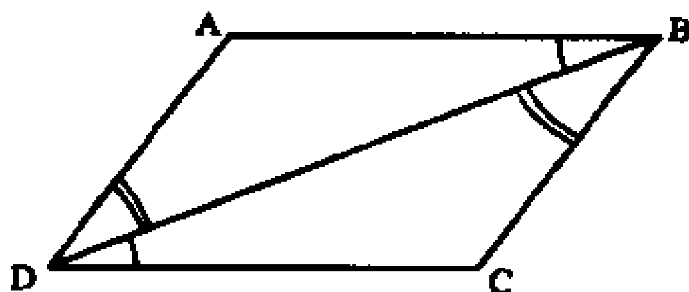


图 8

通过一系列证明把几何学定理加以"整理"的想法肯定很早以前就已出现，后来的评注者也援引过比欧几里得的著作更早的"几何原本"，但它们都未留传下来. 但显然这种不断使假设回 [49] 归的过程不可能无限制继续下去，它必定会碰到不证明的"假设". 众所周知，这正是人们在欧几里得的《几何原本》开头在某些"定义"之后所发现的东西. 现在通常把这些"假设"以几何学

公理的名称收集在一起;我觉得这是一个错误,因为它们并不涉及相同的对象.欧几里得也知道这一点,因为他把它们分为"规定"和"共识"."规定"——也称公设——是平面几何学的性质,而"共识"——到经典时代终了时终于被称为"公理"——则涉及各种类型的"量"(见下面 §5).

4. 几何学——从欧几里得到希尔伯特

正是在欧几里得的《几何原本》里,我们首次看到以柏拉图和亚里士多德的精神按演绎方法构思"数学对象"的性质达到了发达的形态.第Ⅰ至第Ⅵ卷中的"定义"列举了属于平面几何学的对象:点,直线,角,圆,多边形(除三角形和四边形外,只有正多边形得到详细研究).至于这些对象不能为我们的感官所感知这一事实,欧几里得并未留下什么疑问.头两个定义说"点是没有部分的","线只有长度而没有宽度"①.更有特色的是下面两条"规定":

1) 任意两点可由一直线段相连;

2) 直线段可沿两个方向无限延长.

这些性质经常用到,但对物质性的"直线",这两条就是不合理的!

正是从这些"定义"、"规定"和"共识"出发,欧几里得宣布他证明了后面的一系列定理.接下来让我们感到有点惊奇的是,每[50]条定理都附有一张图.或许会认为画这些图的目的是易于领会

---

① 事实上这些只是不能使用的"假定义".一个词的定义必须引进先前已定义的词,而被定义的词则作为简略语.正如帕斯卡采用亚里士多德的教导所表述的,当我们要用到它时,必须"总在心中以所给的定义来代替被定义的东西".但我们从未看到欧几里得用这两个"定义"来代替"点"和"线"这两个词;因此我们可认为他所设想的是,这些词是不予定义的.

定理的证明;据说即使从错误的图出发,几何学艺术仍能很好地进行推理. 但是很快就能看出,其中某些图形起着关键作用,而且那里的论证距印度和中国的几何学家所做的并不很远,后者通常画出图形,而对于证明则只是说由图"看出"①. 例如,欧几里得多次认为下面的断言是当然的(第Ⅲ卷,命题 17;第Ⅵ卷,命题 13):含有位于某一圆内的一个点的直线必与此圆相交;类似地,他也认为下面的断言是当然的(第Ⅰ卷,命题 1 和命题 22):如果一个圆 C 有一个点位于圆 C' 之内,又有一个点位于 C' 之外,则 C 与 C' 相交. 这些事实都不能从他的"规定"中推出. 求助于"可见"对象在第Ⅲ卷命题 8 中甚至更加明显,这里欧几里得研究把圆外的点同圆上的点连接起来的直线段(图 11),并且在该圆上区分"凸圆弧"(相对于外点)和"凹圆弧";对于柏拉图的"绝对图形"而言,他在定义这些概念时本来是会有困难的. 在关于立体几何学的第Ⅺ卷中,他用一条直线或一个半圆绕 [51]

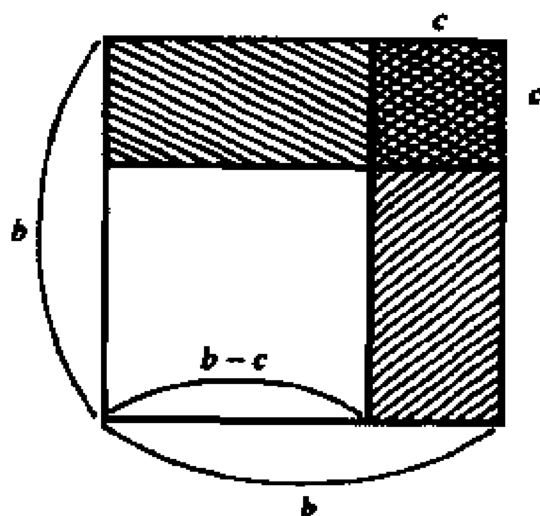


图 9

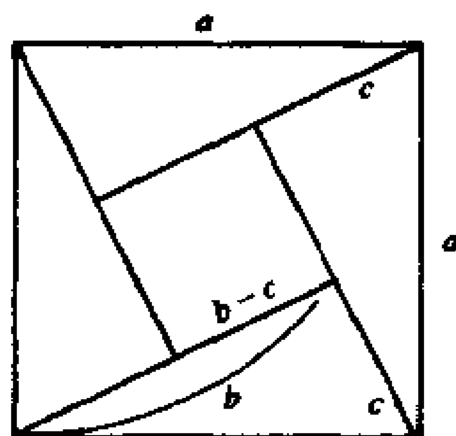


图 10

① 从图 9 能"看出" $(b-c)^2=b^2-2bc+c^2$. 从图 10 能"看出",对于斜边为 a,直角边为 b,c 的直角三角形,有 $a^2=(b-c)^2+2bc$,再据图 9,即有 $a^2=b^2+c^2$,这就"证明"了毕达哥拉斯定理.

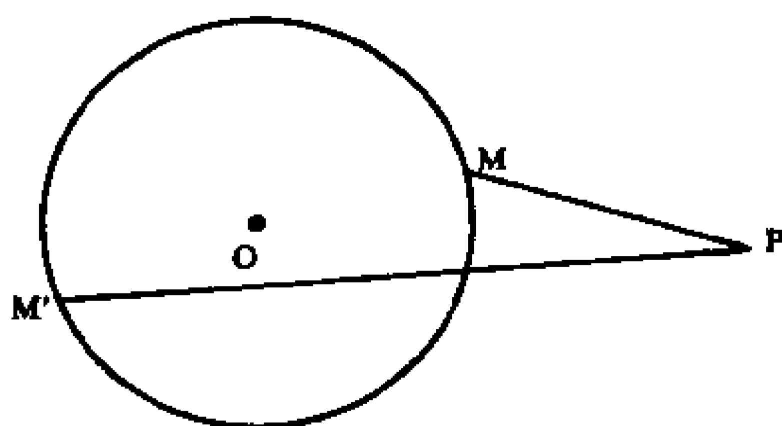


图 11

一"固定"直线"旋转"来"形成曲面"，对这样的定义我们又该说些什么？这样的例子还能举出不少，它们表明了在创立适应于只能"通过思维看到"的对象的本性的词汇表并按这些本性即不用图形来阐述这些对象的性质方面，他本来必须克服的困难.

进一步说，欧几里得考虑的"图形"数目有限，因此他只在少数情形求助于基于研究作图所得到的"根据". 如我们稍后会看到的（§8），古代世界中欧几里得的继承者通过发现和研究新的曲线和曲面，丰富了几何学的领域，从而为文艺复兴后数学的跃进打下了基础. 但如果我们仔细考察，就会看出这种进步伴随着——比欧几里得更甚——借助大量并非系统阐述而是得之于多少精确的作图的性质.

然而，只是由于受到现代公理式理论的严格训练，才能使我们注意到这种不完善性. 除了引起争论的平行公设——欧几里得的第五条"规定"，我们会在第六章中回过来讨论它——外，看来在 16 世纪以前，对这类欧几里得教本或其继承者的论著，并没有很多批评. 似乎长期建立起来的习惯使几何学者迟钝，不能觉察从可感知对象过渡到它们仅是其粗糙形象的对象中的过度自信. 这种觉察在柏拉图和亚里士多德那里表达得如此清楚，以致当我们看到像笛卡儿和帕斯卡这样深刻的思想家——他们毫不犹疑

地向经院哲学发起正面进攻——强有力地宣布几何公理是“明白无误的真理”时，会感到惊讶！而且他们也只是表达了那个时代所有数学家共同的心理状态.在下个世纪，这种情形更加发展，因为高斯和柯西仍然竞相夸耀“几何学的严格性”是其他数学领域的典范.或许这种关于几何对象同其可见形象之间完善和谐的不合理自信对于构造力学和物理过程的数学模型——它们取得了极大的成功——是必要的.只是到了 19 世纪后三十年，随着对于实数系的深入研究，才能考虑把“几何直观”与设想为此种直观提供合理基础的公理系统分隔开来的鸿沟(见第六章).

对欧几里得结构的批评，尤其在 19 世纪导致数学中更高“严格性”的一般运动(我们会在第六章中回到这个论题)中，变得越 [52] 来越多；这些批评的目的不是改正欧几里得在其证明过程中所作的推理，而是纠正欧几里得的推理并没有充分地基于明确陈述的定义和公理这种状况.一般的感觉是，如果你能以恰当的方式使这些推理基础完备，你就能达到完全满意的阐述.当帕施和希尔伯特在 19 世纪末列举了完全明确的公理系统(希尔伯特所列举的是 23 条公理)时，他们所完成的正是这一任务.通过这样的公理系统，欧几里得的全部定理最终都能不用图形加以证明.

像欧几里得那样，希尔伯特从无定义概念开始，但他把这些概念毫无遗漏地列举出来.有三类“本原对象”：点，直线，平面；还有三种“本原关系”：属于(例如一个点属于一直线或一平面)，“位于……之间”(例如在三个点属于同一直线时其中一个点位于另外两个点之间)，“全等”(例如两个线段或两个角[①]全等).

---

① 线段和角是“导出”概念，即这样的词，其定义用到某些公理以及本原概念和本原关系，或用到先前已定义的导出概念：线段 AB 由 A，B 所属的唯一直线 Δ 上位于 A，B 之间的点构成；以 A 为起点并通过点 B 的半直线由线段 AB 以及 Δ 上使得 B 位于 A，C 之间的点 C 组成；角 $\{D_1, D_2\}$ 是具有同一起点的两半直线构成的偶.

立即产生这样的问题：怎样才能既正确地推理又不会定义正在讨论的事物，从而避免定义的无穷回归？答案非常简单：只要迫使人们永不陈述关于几何对象及其间关系的不是制约它们的公理系统(这些公理也毫无遗漏地列举出来)的逻辑推论的任何命题．正如庞加莱所说([9])，可以说这些公理组成了它们所涉及的对象和关系的“隐蔽定义”；后者在某种意义上不出现了，代之以它们的一串“公理式”特性．

在帕施之后，希尔伯特指出了避免作出可能由几何直观提示而不是从公理系统推出的结论的方法：改变几何对象及其间关系的通常名称．希尔伯特提出，我们可以把“点”、“线”和“平[53]面”分别说成“桌子”、“椅子”和“杯子”①．这样一来，希尔伯特列举的前两条公理

1）“两个不同的点属于且仅属于一直线”，

2）“至少有两不同的点属于同一直线”，

就成为

1）“两张不同的桌子属于且仅属子一把椅子”，

2）“至少有两张不同的桌子属于同一把椅子”．

很清楚，不去考虑这些词在日常语言中的含意，上面这样的陈述就不再会有犯无意错误的危险．

这看来像是开玩笑；但事实上这种对于初等几何学所作的意义与名称的分离体现了把数学从它同实在的过于紧密的结合中解放出来的基本过程．它使刚过去的这个世纪中数学取得的出乎意料的胜利及其在物理学中令人惊奇的应用成为可能．我们将在第五章中回过来详细论述这一点．

---

① 任选一个词来规定一个对象，即任选一个词作为刻画该对象性质的陈述的缩略词的可能性，已为柏拉图注意到(信札Ⅶ，343b)．达朗贝尔采纳了这个想法，他在《百科全书》中宣布，没有什么东西不让我们把大家称为“圆”的对象叫做“三角形”．

5. 数和量

当古希腊数学在哲学学派手中精心制作成"假设－演绎"体系①时,古希腊城市日常生活的需要,像别的文明那样,造成了一个职业"计算者"阶层.这些人被称作"计算匠"(logisticians②),除了他们的存在以及柏拉图在《理想国》中(第七卷,525)对他们的轻蔑外,对这些人我们几乎一无所知;柏拉图蔑视他们是因为这些人在分数基础上进行计算,而按照柏拉图的看法,数学只能与自然数有关,"它只能用理性去把握,任何别的方法都不行".

柏拉图这些话无疑是指欧几里得《几何原本》第Ⅶ卷③中涉及任意自然数的一系列令人赞美的定理,其中阐述了自然数的可除性、素数以及自然数分解为素因数之积的初等理论.

数值演算的"计算术"传统只是随同丢番图(约公元4世纪) [54] 才显现出来.他的方法明显地推广了巴比伦泥板上的方法.这是要求出一个或几个"未知"数,它们是方程组的解,我们把每个方程写成具有特定系数的多项式之间的等式形式;这些方程次数最高的可以达到未知数的6次.例如(第Ⅳ卷,19),求3个数,使得这三个数中任意两个之积加1都是一个完全平方.我们可把所述性质写成3个方程:

$$xy+1=u^2, yz+1=v^2, zx+1=w^2.$$

没有必要在这里审察丢番图的解法;他极少求助一般理论.我们要强调的是,对所有问题④,他找的解是自然数或形如 p/q(p, q 是自然数)的分数即我们所称的正有理数.在某些问题中他碰到

① 这一表述应归于皮里(1899).

② 英文译为"logistics"的希腊文词意为实施数值运算.

③ 欧几里得以字母来记自然数,并用线段表示它们.

④ 甚至当所提问题有几个可能的解时,丢番图也几乎总是只考虑一个解.

了两种不可能类型，以下列例子为典型：

$$4 = 4x + 20 \quad (\text{第Ⅴ卷},2),$$

$$3x + 18 = 5x^2 \quad (\text{第Ⅳ卷},31).$$

对于这些情形，丢番图只说所提问题是荒谬的．第二个问题的不可能性来自 41 不是有理数的平方这一事实；我们在 §2 中阐述过$\sqrt{2}$无理性的发现，而前述事实是此发现推广到别的不是平方的自然数情形．

这些不可能性的重要性在于它们都通过创造新的数学对象得到解决——第二种是在古典时代，第一种则在中世纪——，这些新对象与早期毕达哥拉斯学派讨论的对象相比，与明显的物质形象并无紧密联系．

欧几里得及其前辈不需要计算分数．对于他们，分数稍后被代之以两直线段或折线长度之比或两多边形面积之比或两立体体积之比的概念．古希腊人把长度、面积和体积包括在“量”这个一般概念之中（量又有 3 种不同类型），而欧几里得的“共识”则可看作关于这个概念的“希尔伯特式”公理式处理的肇始．事实上，欧几里得既没有定义量这个概念，也没有定义同类量之间的[55]两种“本原”关系即“大于”（我们把它写为 $A > B$）与“两者之和”（我们把它写为 $C = A + B$）．他所做的是列举（但并非毫无遗漏）联系这些概念的某些性质；例如，他的第四条“共识”用我们的记号就可写为：

如果 $A > B$，则 $A + C > B + C$．

量 A 的整数倍 pA 通过把量 A 加自身 p 次得到．欧几里得然后称两个同类量 A 与 B 是可公度的，如果存在第三个同类量 C，使对两个自然数 p, q 有 $A = pC$，$B = qC$；他然后称（第Ⅹ卷，命题 5）A 与 B 同自然数 p 与 q 具有“同比”（我们当然不管欧几里得，把所述事实写成 $A/B = p/q$）．

这样，柏拉图学派的希腊数学家，不像丢番图那样，说正方形的对角线与其边没有“比”，而是宣称两个同类量总有比，即使

它们不可公度也是如此．他们然后通过大胆的构造（直到19世纪人们才充分理解这种构造），从用于处理自然数之间比的不等式和加法概念出发，成功地对这些一般的"比"定义了不等式和加法概念．我们不在这里阐述这一令人注目的发明（或许应归之于那个时代最伟大的数学家尼多斯的欧多克索斯），它与现代概念十分接近（见附录1）．

可惜，由于古希腊人对于比的乘积的概念，他们未能使这些一般的"比"的演算体系成为合用的形式．事实上，当欧几里得设想长度之间的两个比 A/B 与 C/D 的乘积时（《几何原本》第Ⅵ卷，命题23），他表明这是两个矩形面积之间的比 $(A\times C)/(B\times D)$．此外，欧几里得没有定义任意两个比之和，而在他所考虑的特殊情形，求和的比具有相同的分母（《几何原本》第Ⅴ卷，命题24）．于是在这种体系下，丢番图（以及甚至§1中所引的他的巴比伦前辈）的演算就行不通．

然而，如果我们不去理睬哲学偏见，就不难把两种观点统一起来而仍然忠实于欧几里得的框架．在重新进行研究之后，首先在穆斯林世界，然后在西方世界，有几位数学家各自独立地注意到这种可能性：11世纪的诗人兼数学家奥马·海亚姆，16世纪的意大利数学家R·邦贝利，最后是笛卡儿，他的声望使这一改革得到最终确认．这一革新在于只考虑长度之间的比，而且把这样的比都简化为 OX/OU 的形式，其中 OU 是一个一次性地选定的线段，而 X 是延长 OU 所得半直线 Δ^+ 上的任意的点．根据欧 [56] 几里得断言"第四比例项"存在性的命题（《几何原本》，第Ⅵ卷，命题12），这种简化是可能的．现在可以不再老是谈论"比"，而简单地说考虑 Δ^+ 的点 X，而我们必须做的就是定义两个点之间的关系 $X<Y$ 以及运算 $X+Y$ 和 XY，后两者也应定义为点，而不像欧几里得那样定义为别种类型的对象．

$X<Y$ 的定义是显见的：X 必须位于 O 与 Y 之间且与 Y 不同．对于 $X+Y$ 和 XY，有许多处理方法．在我看来，对希尔伯特

采用的方法加以某些修改是最简单的，它只需要作平行线，并只须基于欧几里得涉及全等或相似三角形的定理.

考虑不是 Δ^+ 的延长的另一半直线 Δ' 以及 Δ' 上的点 A. 图 12 和图 13 表明了 $Z = X + Y$ 和 $W = XY$ 的构造(当然它们可以不用图形加以描述). $Z' = Y + X$ 和 $W' = YX$ 的构造用虚线标记. 这证明了关于 Δ^+ 上点的运算的两条基本性质:

$$Y + X = X + Y, YX = XY.$$

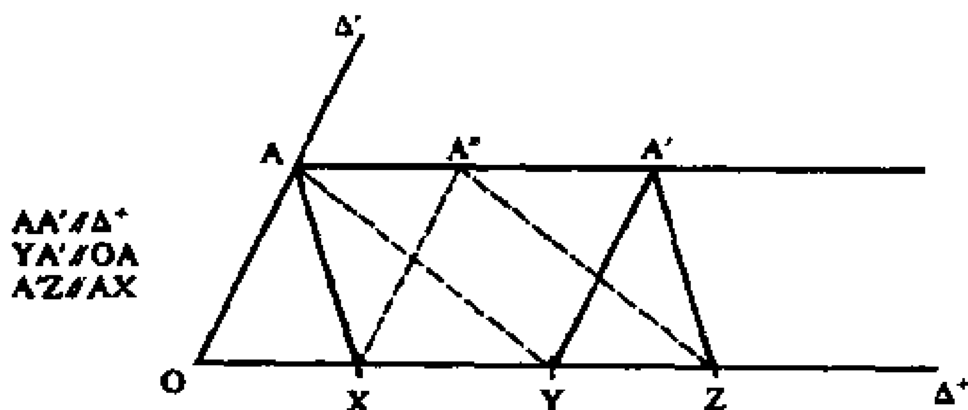
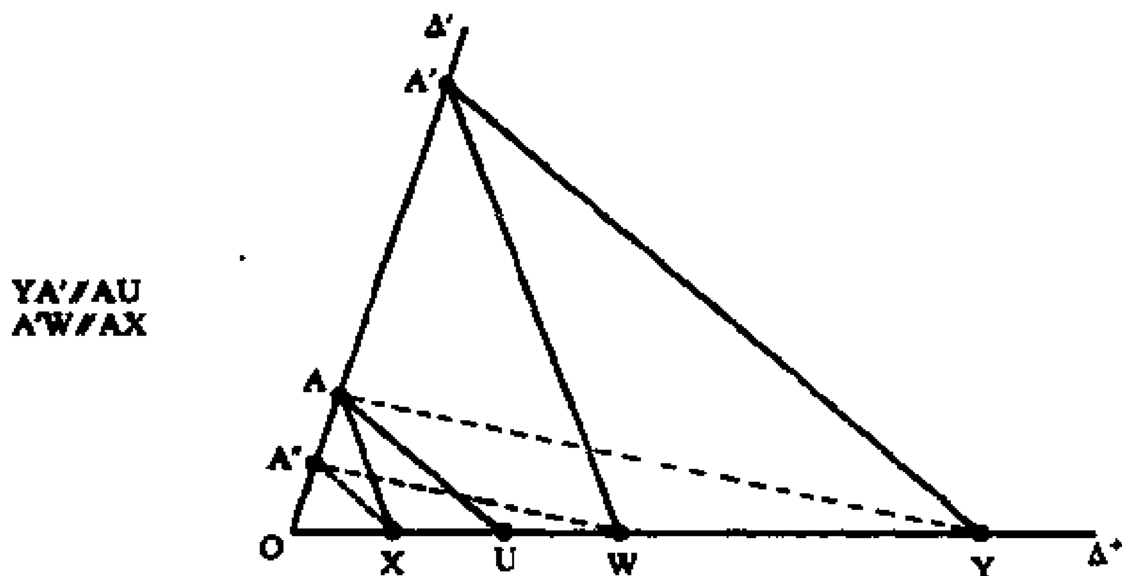


图 12



[57]

图 13

图 14 表明,对 Δ^+ 的所有点 X,存在其“逆”点 X',使得

$$XX' = X'X = U;$$

X' 也可写为 $1/X$ 或 X^{-1}(由于显见的方程 $UX = XU = X$, U 也可写为 1).类似的推理可证明附录 2 中列举的除最后两条之外的全部性质.

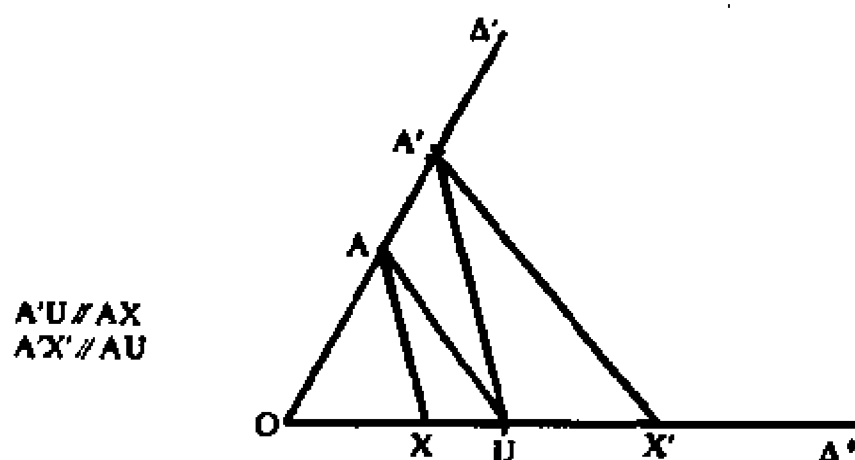


图 14

要注意,在上面的表述中,不必区分可公度的比与不可公度的比.对这些构造直接推广一下也能破除丢番图碰到的第一种不可能性.正是这个问题导致公元 4 世纪的印度数学家引进“负数”.负数的使用传播得很慢,到 18 世纪末仍然遇到阻力.事实上,需要做的全部事情就是把上面的构造拓广到包含 Δ^+ 的直线 Δ 上所有的点,而不只是考虑 Δ^+ 上的点.这样方程 $X + Y = X$ 总有解,即使 $X < Y$ 时也是如此(图 15).我们把这个解写为 $X - Y$.特别地,$0 - X$ 写为 $-X$,这给出 $X + (-X) = 0$. [58]

现今我们说由这种构造给出的 Δ 的点是实数,Δ^+ 上的为正,Δ^- 上的为负.注意这种构造立即“解释”了“正负号规则”

$$(-X)Y = -(XY),$$

以及特别地两个负数之积为正数这个事实,此点对于非数学家看来似乎难以理解(图 16).

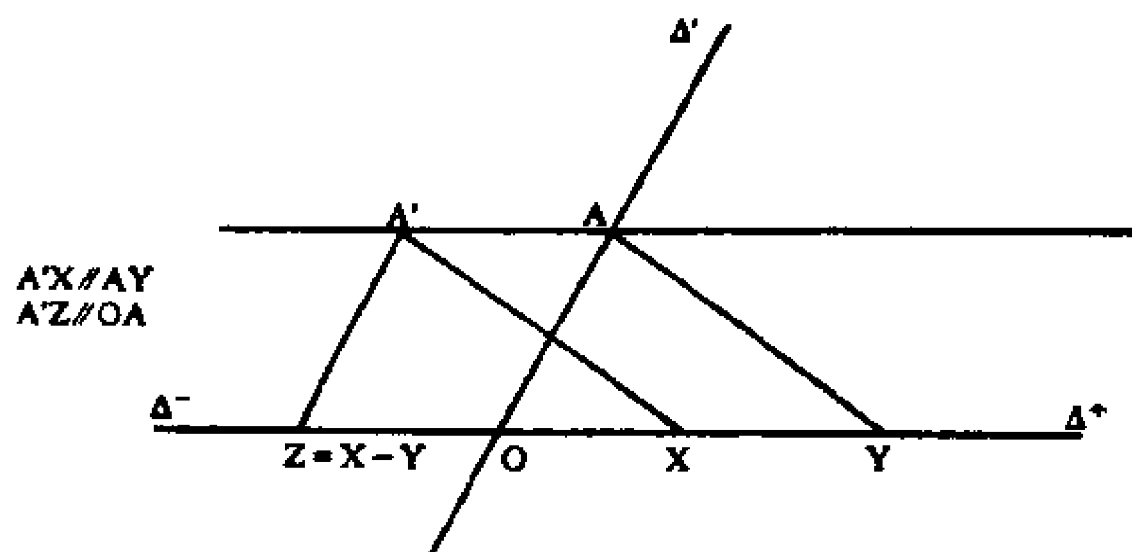


图 15

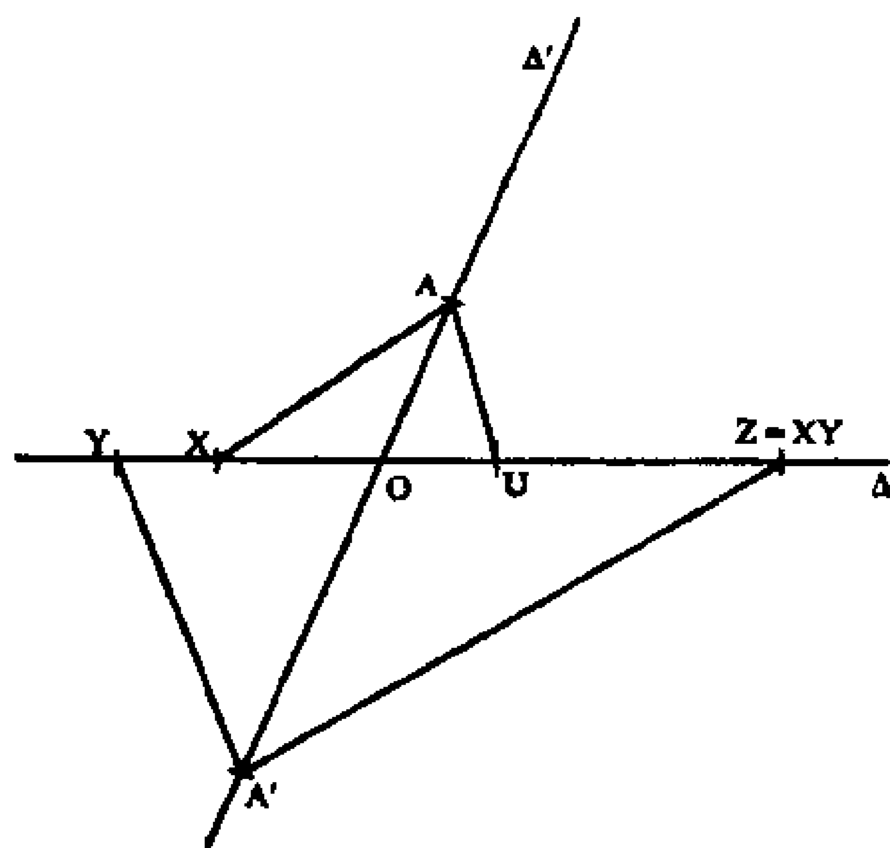


图 16

6. 逼近的想法

在应用数学时,几乎总是用一个有理数 X 代替实数 Y,X 称为 Y 的近似值;这就是说,差 $Y-X$ 介于 $-\varepsilon$ 与 ε 之间,此处 ε 是一正数,称为用 X 近似 Y 的**误差限**. X 称为 Y 的**误差限为** ε 的近似值,而 Y 本身则看作"精确值". 巴比伦人已经知道怎样求$\sqrt{2}$的有理近似值,而古希腊人则发明了求得所有平方根$\sqrt{a}$(a 是自然数)的具有任意固定误差限的近似值的一般方法.

这样的近似值的存在性来自由阿基米德明确表述的一条一 [59] 般公理[①];他常用到它,因而通常称为**阿基米德公理**. 此公理说,如果 Y, X 是两个正数,则总存在自然数 m,使得 $Y<mX$. 于是,如果 p 是使得 $Y<(p+1)X$ 的最小自然数,对我们有 $pX\leqslant Y<(p+1)X$,因而 pX 以误差限 X 近似于 Y. 阿基米德公理的一个等价形式是总存在自然数 m,使得

$$\frac{1}{m}Y<X^{②}.$$

我们可顺便注意这样的公理与感觉经验相距多远:除非 X 和 Y 都既不过大也不过小,否则就不可能作出把等于 X 的线段不断首尾相接直到超过 Y 为止这种操作的有形形象.

由于 X 能再分为任意个数的相等部分(《几何原本》,第Ⅵ卷,命题 10),一旦得到近似值 pX,我们可对 $0\leqslant m\leqslant 9$ 作

$$pX+m\frac{X}{10}.$$

如果 $q+1$ 是满足

$$Y<pX+(q+1)\frac{X}{10}$$

① 欧几里得在《几何原本》第Ⅹ卷命题 1 中用过它.

② 这样,不存在正数 Z,它小于所有其余的正数(即没有"无穷小"数).

的最小的数，则

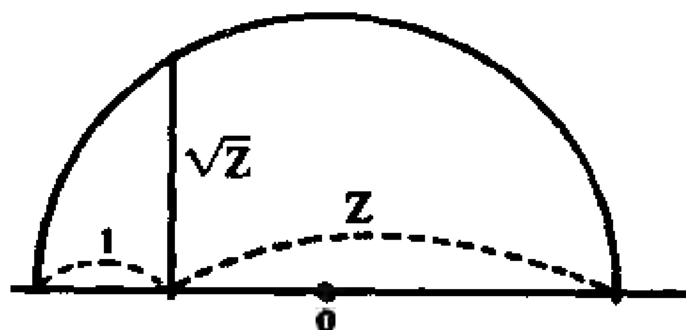
$$pX + q\frac{X}{10}$$

是 Y 的误差限为 $X/10$ 的近似值.我们可继续下去得到

$$pX + q\frac{X}{10} + r\frac{X}{100},$$

并不断往下进行.

这样的过程给出 Y 的误差限为任意小的近似值，因为对任一数 $Z>0$，存在数 n，使得 $X/10^n<Z$. 然而，这里假定这种过程应用于一个已知的"精确的"数 X. 但当 $Z>0$ 是一已知数时，对 $X=\sqrt{Z}$ 也能这样说吗？欧几里得给出的关于"比例中项"的作图(《几何原本》第Ⅵ卷，命题 13)(图 17)通过半圆和直线相交回答了这个问题.



[60] 图 17

从公元前 5 世纪起，古希腊人还面临关于我们所称的"立方根"的类似问题.我们将在 §8 中看到他们对这个问题怎样"作图"，以及他们的几何概念怎样阻碍他们对"n 次根"($n\geq 4$)进行类似的作图.

在丢番图之后，这些顾虑几乎不再出现. 14 世纪的奥雷姆暗地承认，对于每个自然数 n 和每个实数 $X>0$，存在数 $Z>0$，满足 $Z^n=X$. 我们把这个数写为 $X^{1/n}$. 他甚至由此导出分数指数幂 $X^{p/q}$ 的想法，甚至还有设想无理数指数幂诸如 $X^{\sqrt{2}}$ 这样的大胆想法. 在 16 世纪 S·斯蒂文的工作中，出现了推广上述近似

过程的一种逼近方法，它同时还能看作一种存在性证明．他讨论一个方程 $P(t)=0$，其中 $P(t)$是多项式，我们写为

$$P(t)=a_0t^n+a_1t^{n-1}+\cdots+a_n,$$

而 $a_0>0, a_n<0$. 由此有 $P(0)<0$，易于证明存在自然数 A 大到足以使 $P(A)>0$（见附录 3）．斯蒂文依次以 $1,2,3,\cdots$代入$P(t)$中的变量 t，到使得 $P(q)<0$ 的最后的自然数 q 为止，于是有 $P(q+1)\geqslant 0$. 然后他以数

$$q+\frac{1}{10},q+\frac{2}{10},\cdots,q+\frac{9}{10}$$

代入 t，直到使得

$$P\left(q+\frac{q_1}{10}\right)<0$$

的最后的自然数 $q_1\leqslant 9$ 为止．他以数

$$q+\frac{q_1}{10}+\frac{1}{100},\cdots,q+\frac{q_1}{10}+\frac{9}{100}$$

进行同样的操作，并不断这样做．这样他就得到一个能无限延续的区间套(图 18)

$$[b_0,c_0],[b_1,c_1],\cdots,[b_k,c_k],\cdots \tag{1}$$

而其长度

$$c_k-b_k=\frac{1}{10^k}.$$

加之我们还有 $P(b_k)<0, P(c_k)\geqslant 0$.

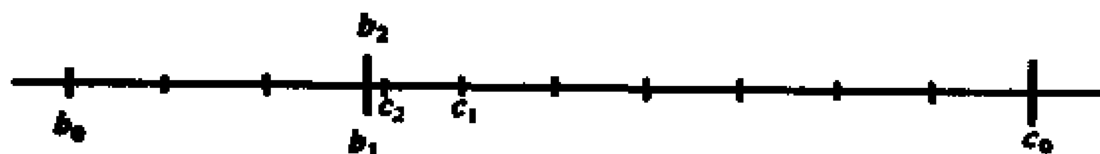


图 18

他然后隐含地假定存在含于所有上述区间中的数 Z. 此外，

不会存在两个这样的数 $Z' < Z$，因为若如此就会对每个自然数 k 有 $0 < Z' - Z < 1/10^k$，这与阿基米德公理矛盾. 下一步他假定有 $P(Z) = 0$，此点可用初等方法证明(见附录 3). 这一程序还给 [61] 出得到 Z 的具有任意小误差限的近似值 $b_0, b_1, \cdots, b_k, \cdots$ 的常规方法. 把这方法用于多项式 $P(t) = t^n - X$ 就给出逼近 $X^{1/n}$ 的一个程序.

以后几个世纪的数学家使用并推广了这种方法；不过一直等到柯西才从中清楚地提炼出区间套公理，此公理断言含于序列(1)的所有区间中的实数的存在性，(1)中每个区间包含后一个区间①. 在实数系的公理式理论中，这个公理是附录 2 中列举的最后一条公理；无疑它也是最重要的公理，因为它是控制整个分析运转的关键.

7. 代数学的演进

有可能对未知数进行演算，就像它们是已知数那样，这件事并非显然，它常常是青少年难于理解的事情. 这是代数学的真正起点，如我们已说过的，它是已为公元前第二个一千年中巴比伦书记人员所跨过的一道门槛. 可是，虽然几何学达到基本上最后确定的形式所花时间没有超过一个世纪，但对于代数学，则要等到丢番图以后 13 个世纪之后，它才变为我们今天所知的形式.

阻碍代数学发展的主要障碍在于代数学家使用日常语言描述代数中用和与积这两种基本运算进行演算的方式. 例如，下面是花拉子米(9 世纪)对求解二次方程的描述，花拉子米在中世纪是如此著名，以至其姓氏稍有讹用而成为我们今天使用的"algorithm"(算法)这个词. 花拉子米的描述如下：

问题. 当我们把一个正方形面积加上其一边长度之十倍等于 39 时，此正方形必是什么？(用现代记号，就是 $x^2 + 10x = 39$.)

---

① 我们记得这些区间都包括其两个端点.

解. 把所加边长的倍数除以 2，得 5. 把该数自乘，得乘积 25. 把此数与 39 相加，得 64. 取此数的平方根即得 8，从该数中减去边长倍数之半，剩下 3. 此即所求正方形的边长，因而所求正方形面积等于 9.

容易想象，当所提问题稍微复杂一些时，用这种方式写出的解的描述必定会变得像法律文件那样含糊难解. [62]

这个困难非常缓慢地导致这样的认识：为使一连串运算易于理解，需要一种速记法；这里我们碰到记号问题，每当引进新对象后总会出现这个问题，或许它会不断折磨数学家们.

毫无疑问，丢番图是最早使用简略记号的代数学家. 在涉及一个未知量的问题中，他用各种专门符号来表示此未知量的前六个幂. 中世纪的代数学者仿效他的做法，采用如下写法：

co 或 *c*(cosa)	未知量 x
ce(censo)	平方 x^2
cu(cubo)	立方 x^3
cece(censo di censo)	幂 x^4

然而，当问题涉及另一个未知量时，丢番图及其绝大多数继承人并没有第二个未知量的记号. 15 世纪末，卢卡·帕奇奥利把“加”和“减”写为 p 和 m①，并对 2 次，3 次和 4 次根用 R 或 $R2$，$R3$，$R4$ 或 RR 表示. 例如，对于

$$\sqrt{.40-\sqrt{320}},$$

他写为

$$RV40\tilde{m}R320$$

(V 表示对其后的表达式用括号括起；在 R·邦贝利的论著中，我们发现用 L 和 ⅃ 表示括号). 记号 + 和 − 只是到大约 1480 年时才出现，记号 = 出现于 1557 年，而关于不等的记号 < 和 > 则出

① 分别对“piu”和“meno”用的缩写. ——英译本注

现于 1631 年.

通向符号化的关键一步归功于韦达.在不必对给定的量给出数值的问题中,他不但用字母标记未知量,也用字母标记给定的量.这就使得演算的描述大大缩短.例如,我们写为 $bx^2+dx=z$ 的式子,韦达写为

B in *A* Quadratum, plus *D* plano in *A*, equari *Z* solido①.(*B* 乘以 *A* 平方,加面积 *D* 乘以 *A*,等于体积 *Z*.)

指数(正的或负的)记号已为 15 世纪的 N·许凯和 M·斯蒂费尔所使用,然后在 16 世纪又为斯蒂文所使用(但韦达没有用指数记号).随着笛卡儿的论著,它成为流行的用法.最后还剩下对于一列数,当我们并不需要特别指明它有多少项时,引进下标记号.这个想法出现于牛顿和莱布尼茨的论著中,但直到 19 世纪末,许多数学家还是喜欢写

$$a, b, c, \cdots, l$$

而不用现代记号

[63] $$a_0, a_1, \cdots, a_n,$$

后者能使我们对 $n+1$ 项进行演算而毫无不便之处.只是到了这段很长的演化进程的末尾,才有可能写出"一般多项式"

$$a_0x^n + a_1x^{n-1} + \cdots + a_n = \sum_{j=0}^{n} a_j x^{n-j},$$

这就有可能表述和证明代数学中的一般定理.

8. 坐标方法

众所周知,我们把 17 世纪数学三项重大创新之一——在几何学中引进坐标方法(也称为"把代数应用于几何",后来称为"解析几何")归功于笛卡儿(同时归功于费马).不过笛卡儿和费

① 忠于柏拉图传统的韦达只对同类量相加,因之 *B* 和 *A* 是长度,*D* 是面积,*Z* 是体积.

马并非没有先驱者，因为由方程定义平面曲线的最早例子可追溯到公元前4世纪．我们已看到（§2），作面积为另一正方形面积两倍的正方形是容易的．类似的“倍立方体”问题无疑在相当早时即已提出，但看来它对古希腊数学家要困难得多．倍正方形问题相当于作一长度 x，使得它是两个长度 a 与 $2a$ 的“比例中项”，即使得

$$\frac{2a}{x}=\frac{x}{a};$$

因为由两个比乘积的定义（§5），有

$$\frac{x^2}{a^2}=\frac{2a\times x}{x\times a}=2.$$

正是这种处置倍正方体问题的方法导致希俄斯的希波克拉底（公元前5世纪）把倍立方体问题归结为作 a 与 $2a$ 之间的两个“比例中项”即作两个长度 x,y，满足方程

$$\frac{2a}{x}=\frac{x}{y}=\frac{y}{a}. \tag{2}$$

事实上，对于古希腊人，3个长度 A,B,C 的乘积是边长为 A,B,C 的平行六面体的体积，从而关系(2)导出（欧几里得，《几何原本》第Ⅷ卷，关于可公度比的命题12）

$$\frac{y^3}{a^3}=\frac{2a\times x\times y}{x\times y\times a}=2.$$

[64]

看来完成这个作图一点也不比原先的作图容易．但欧多克索斯的学生梅内克缪斯注意到，(2)也能写成两个方程联立的形式：

$$x^2=2ay, \tag{3}$$

$$xy=2a^2. \tag{4}$$

梅内克缪斯然后产生这样的想法：在平面上作两条互相垂直的半直线 OX 和 OY，并在 OX 上线段 $x=OP$ 与 OY 上线段 $y=OQ$ 之间分别考虑关系式(3)，(4)．但对 OY 上每个点 $Q(OQ=y)$，在过点 Q 作的平行于 OX 的直线（图19）上存在点 M_1，使得 $x_1=OP_1=QM_1$ 满足 $x_1^2=2ay$（$2a$ 与 y 之间的“比例中项”），

图 19

也存在点M_2，使得 $x_2 = OP_2 = QM_2$ 满足 $x_2 y = 2a^2$($2a$，y 与 a 之间的"第四比例项")。

当 Q 在 OY 上变动时，点 M_1，M_2 描绘出两条曲线 C_1，C_2. 梅内克缪斯稍后发现，这两条曲线可以作为平面与圆锥面的截线得到. 在欧几里得和阿基米德之后，阿波罗尼奥斯(公元前 3 世纪)深入研究了这些称为圆锥曲线的平面曲线，他称 C_1 为抛[65]物线，称 C_2 为双曲线.

如果 N 是这两条曲线的交点，则线段 $OH = NK = x$ 和线段 $OK = NH = y$ 给出(2)的解. 稗官野史说柏拉图不喜欢这个解法，因为它用到直线和圆以外的曲线；不过这类通过曲线相交的

“作图”在欧几里得继承者中多少还是流行的(即使它确实暗示一些“可以看出”但没有证明的性质),而且笛卡儿和牛顿也仍然给以高度评价[1].

笛卡儿和费马所用方法的重要性在于这样的事实:它使平面几何学中任何问题翻译为代数学中与之等价的问题成为可能.这是因为,按照下面的“词典”,希尔伯特公理系统中的对象和关系可以解释为实数论中的对象和关系:

点 M	数偶(x,y)[2] 我们写为 $M=(x,y)$
直线	方程 $ax+by+c=0(a^2+b^2\neq 0)$
属于……的点	满足……的数偶
$A=(x_1,y_1)$与 $B=(x_2,y_2)$之间的点	数偶$(tx_1+(1-t)x_2, ty_1+(1-t)y_2)$, 其中 $0<t<1$
线段 AB 的长度	$((x_1-x_2)^2+(y_1-y_2)^2)^{1/2}$
圆	方程 $x^2+y^2+ax+by+c=0$ $(4c<a^2+b^2)$

最后,对于直线 AB,AC 之间的夹角 α,有公式

$$\cos\alpha=\frac{(x_1-x_0)(x_2-x_0)+(y_1-y_0)(y_2-y_0)}{((x_1-x_0)^2+(y_1-y_0)^2)^{1/2}((x_2-x_0)^2+(y_2-y_0)^2)^{1/2}},$$

这里 $A=(x_0,y_0)$,$B=(x_1,y_1)$,$C=(x_2,y_2)$.

通过用3个数的数组(x,y,z)代替点,用方程 $ax+by+cz+d=0(a^2+b^2+c^2\neq 0)$代替平面等等,也能类似地对“立体几何学”加以“翻译”.

到了19世纪,当数学基础终于更深入地建立后,就说初等

---

[1] 可以用同样的方式定义 $2a$ 与 a 之间三个、四个、……“比例中项”,但古希腊人不能进行3次以上根的“作图”,因为在他们看来,多于3个长度的乘积就会没有意义.

[2] 当然$(x,y)\neq(y,x)$,除非 $x=y$.

[66] 几何有一个实数论中的模型.这是把数学中并不相似的两个部分联结起来的一系列“桥梁”的第一座.我们还会在第六章中遇到某些这种桥梁,它们是数学思想基本统一性的越来越鲜明的佐证.

再进一步,平面上的直线、圆和圆锥曲线可通过形如 $P(x,y)=0$(P 是实系数的一次或二次多项式①)的方程定义这个事实自然导致数学家研究同类但对次数不加限制的方程所定义的曲线.这是数学中一个新分支——代数几何学的滥觞,它在古希腊人已知的曲线清单上添加了许许多多新曲线.在300年的研究并取得大量成果之后,它如今仍是数学中最活跃的分支之一.

坐标方法也是17世纪另外两个伟大进展——引进函数概念和微积分的基础.人们常说,古希腊人的数学概念基本上是静态的,与之相反,变化观念主导了现代科学思想.当然,欧几里得的《几何原本》集中研究位置和大小都固定的图形.然而,自从古希腊思想肇始以来,从形式上或从本性上理解运动和变化的尝试一直萦绕在哲学家心头,而从知道如何测量时间后,直线或圆周匀速运动概念就已得到清楚的考察.我们知道,在古希腊人的天文系统中,他们正是试图用这些运动的合成说明行星的轨道.虽然时间概念没有包含在古希腊几何学中②,但至少有两条平面曲线——希庇阿斯割圆曲线和阿基米德螺线是通过匀速运动的合成来定义的③.

---

① x,y 的多项式是一些单项式 $a_{jk}x^jy^k$ 之和;其次数是 $j+k$ 中最大的数.

② 或许我们应当把埃利亚的芝诺的著名悖论看作试图把时间和运动引进“假设-演绎”系统的失败结果.

③ 希庇阿斯割圆曲线由如下的点 M 所描绘:OM 以匀速绕 O 转动,而 M 在 OY 上的投影 Q 在 OY 上匀速运动(图20).阿基米德螺线由如下的点 M 所描绘:D 以匀速绕 O 转动而 M 沿 OD 匀速运动(图21).

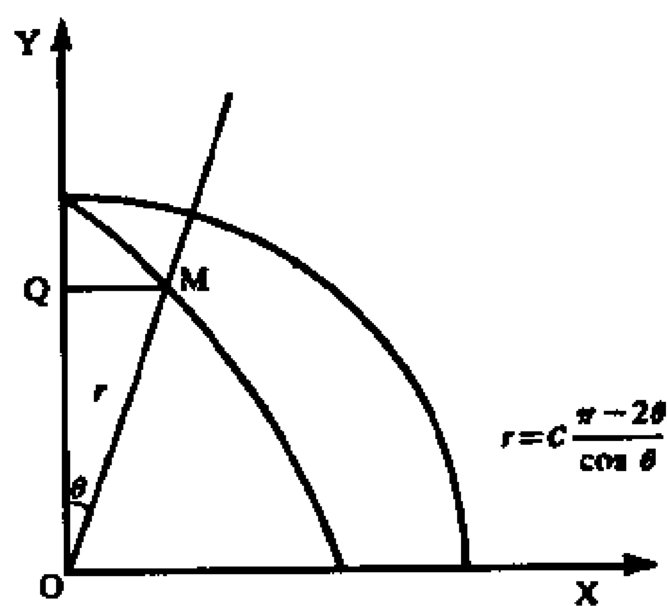


图 20

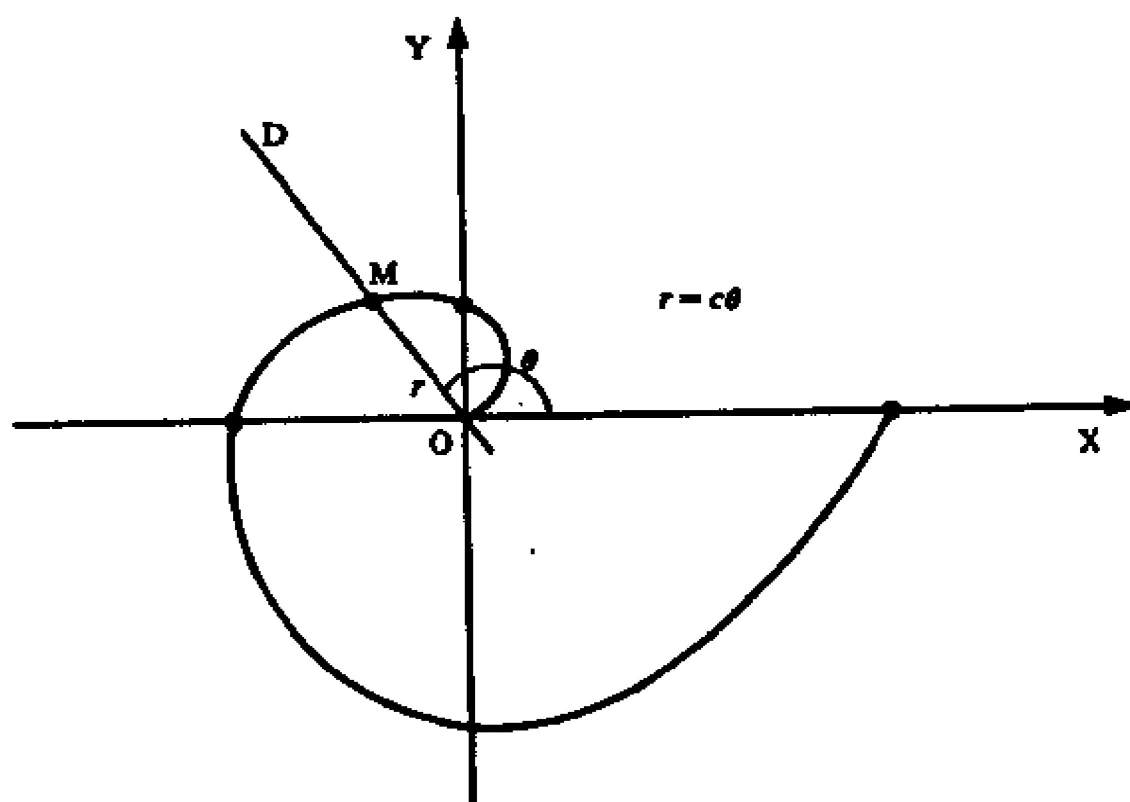


图 21

[67] 看来主要是对于直线但不一定是匀速的运动——尤其是落体运动，这是中世纪哲学学派考虑的一个突出主题——的研究使得 14 世纪的奥雷姆产生了——可能在历史上首次——通过图象来表示量在时间进程中的变化的想法，这里图象上点的横坐标表示时间而纵坐标则表示变化的量在该时刻的值（图 22）；这样得到的点构成了所研究的图象．奥雷姆还认为，可以用任何
[68] 能用数表示的“量”①代替时间作为横坐标．这样的做法现在已无处不在，而且常常误用．

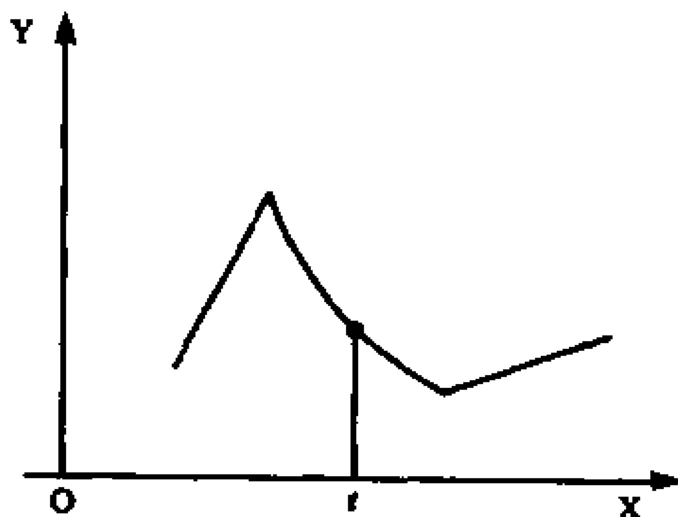


图 22

在 17 世纪，上述想法与坐标方法联结起来，使得数 y“依赖于”在区间 I 上变动的数 x 这个想法为人们所熟悉．到 17 世纪末，就说这样的 y 是 x 的函数．它的图象是一条曲线，并与过 I 的每个点且平行于 OY 的直线只有一个交点．反之，每条具有前述性质的曲线定义一个函数；例如，圆心为 O 半径为 1 的位于

① 它不一定是古希腊人理解的“量”．一个例子是以摄氏度测定的温度：一个温度小于另一温度这概念具有物理意义，但其度量是另两个温度之“和”的温度却没有物理意义．

OX 上方的半圆(图 17)在区间 $-1 \leqslant x \leqslant 1$ 上定义函数

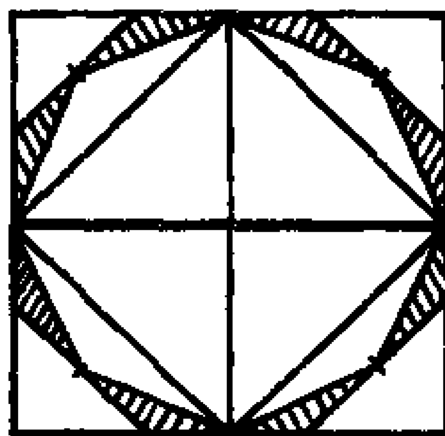
$$y=\sqrt{1-x^2}.$$

正是这种对应使得在 19 世纪作为柏拉图意义下的数学对象定义函数一般概念成为可能(见第五章, §3,B).不过直到那时仍然缺乏对于理论基础的关注.然而尽管函数概念可能很"直观",它却开辟了一个人们未曾预料到的飞速进展的时代,这种进展在数学应用中决不比数学中更慢,所有数学家都竞相投入这一进步之中.这首先是由于这样的事实:函数概念是 17 世纪第三项数学发明,或许是整个数学史中最重要的发明——微积分的基础.

[69]

9. 极限概念与微积分

留传至今通过"区间套"过程(§6)确定一个数的最早例子是由欧几里得给出的(《几何原本》第Ⅻ卷,命题 2)对于圆面积的逼近.他在圆内作内接正方形 P_1,然后依次取前次多边形各边所对弧的中点,得到 8 边,16 边,…,2^{n+1}边…正多边形 P_2, $P_3,\cdots,P_n,\cdots$.他同时考虑各边在前述多边形顶点处与圆相切的外切正多边形 $Q_1,Q_2,\cdots,Q_n,\cdots$(图 23).设 p_n 是P_n 的面积,



(画有影线部分面积为 q_2-p_2)

图 23

q_n 是 Q_n 的面积,则这些数构成一个区间套序列

$$[p_1,q_1],[p_2,q_2],\cdots,[p_n,q_n],\cdots$$

的各端点.欧几里得通过优美的几何论证证明

$$q_n-p_n\leqslant\frac{1}{2}(q_{n-1}-p_{n-1}).$$

这样他就能断言,含于所有这些区间中的唯一的数测量出"该圆的面积"(见附录4).

这一构造是所谓"穷竭法"的简单例子,此法或许是欧多克索斯发明的.无论如何,正是欧多克索斯应用此法证明圆锥的体积是同底同高圆柱体积的三分之一(《几何原本》第Ⅻ卷,命题10).在阿基米德手中,穷竭法产生了一大批新结果:抛物线形的面积,球面面积和球的体积,旋转抛物体的体积,等等.

17世纪中坐标的应用使得推广穷竭法成为可能.例如,如果 $y=f(x)$在区间

$$I:a\leqslant x\leqslant b$$

上是正递增函数,则我们能用这种方法求此函数的图象与 OX 轴之间所围面积 S 的值.把 I 相继分为 $2,4,8,\cdots,2^n,\cdots$个相等部分,对每个子区间$[t_k,t_{k+1}]$考虑以此区间为底、位于所给曲线之下的矩形的面积 [70]

$$\frac{(b-a)}{2^n}\cdot f(t_k)$$

和以此区间为底,位于所给曲线之上的矩形的面积

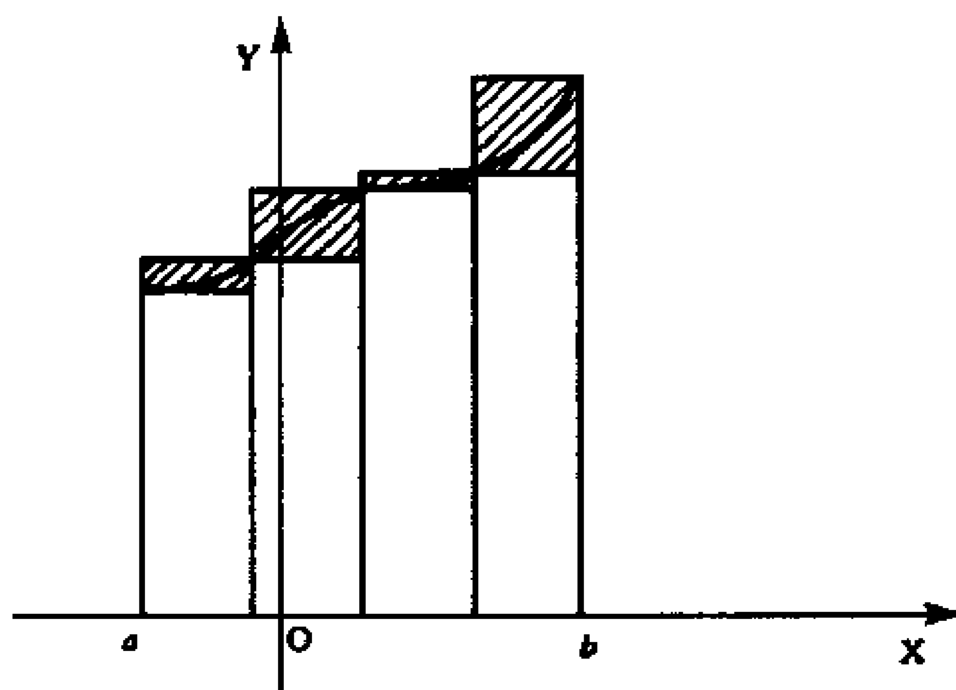
$$\frac{(b-a)}{2^n}\cdot f(t_{k+1}).$$

如果 S_n'是前一种矩形面积之和,S_n''是后一种矩形面积之和,则我们有 $S_n'\leqslant S\leqslant S_n''$,且有

$$S_n''-S_n'=\frac{(b-a)}{2^n}(f(b)-f(a))$$

(图24).

这样数 S 就定义为含于区间套中所有区间$[S_n',S_n'']$中的



(画有影线部分面积为 $S_2'' - S_2'$)

图 24

唯一的数.这是费马于 1636 年后关于曲线 $y = x^m$(m 是大于 1 的自然数)所用的方法.他通过一个简单的代数公式明显地给出 S_n'和 S_n'',从而得到公式

$$S = \frac{b^{m+1} - a^{m+1}}{m+1}$$

(见附录 4).

后来,费马通过更好地选择把 I 划分为越来越小区间的方法,把上述结论推广到 $m = p/q$ 为任何(不等于 -1 的)分数的情形.

这个方法(连同些许可取较少子区间的改进)至今仍是用计算机计算面积的基础之一.要不是同样的数学家们同时向另一个在古典时代只是偶尔碰到的看起来很不相同的问题,即确定平面曲线的切线问题发起冲击,或许前述方法本来不会引起那么大的兴趣. [71]

对于古希腊人,曲线 C 在其上一点 M 处的切线是过 M 的一条直线 D,使得 C 上与 M 邻近的点都位于 D 的同一侧(图25):在点 M 处 D"切触"所给曲线.欧几里得证明,以 O 为中心的圆上点 M 处的切线垂直于 OM(《几何原本》,第Ⅲ卷,命题16);而把圆锥曲线描绘为圆锥面与平面相截得到的曲线使得求出这些曲线的切线成为可能.然而,在古典时代,除圆锥曲线外,只知道一条曲线即阿基米德螺线(§8)的切线的确定方法,而且我们也不知道他是怎样猜出这些切线的作图方法的[①].

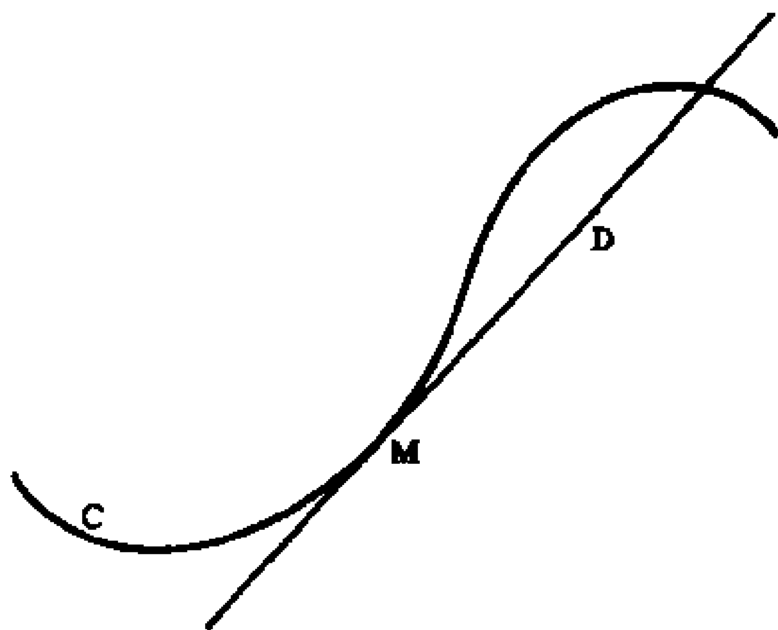


图 25

正是费马又于1636年对于曲线 $y=x^m$(m 是大于或等于2的自然数)用坐标方法系统地向这个问题发起冲击.问题在于由方程 $Y=aX+b$ 给出的通过点(x,y)(其中 $y=x^m$)的直线是否在该点处"切触"曲线 $y=x^m$.这样,必须知道对于 OX 上横坐标

① 阿基米德进行了这些切线的作图,但未加任何解释,然后用"反证法"(附录1)证明此直线在触点处不可能穿过所给的曲线.

为 $x+h$（$|h|$充分小）的点 P，有向线段$\overline{RS}$的正负号（图 26）. 由于 D 过点 M，故有 $x^m = ax+b$，从而由二项式公式，

$$\begin{aligned}\overline{RS} &= (x+h)^m - (a(x+h)+b) \\ &= mhx^{m-1} + h^2Q(x,h) - ah \\ &= h(mx^{m-1} - a + hQ(x,h)),\end{aligned}$$

其中 $Q(x,h)$是 x,h 的多项式，而且存在数 A，使对$|h|\leqslant 1$有$|Q(x,h)|\leqslant A$（附录 3）. [72]

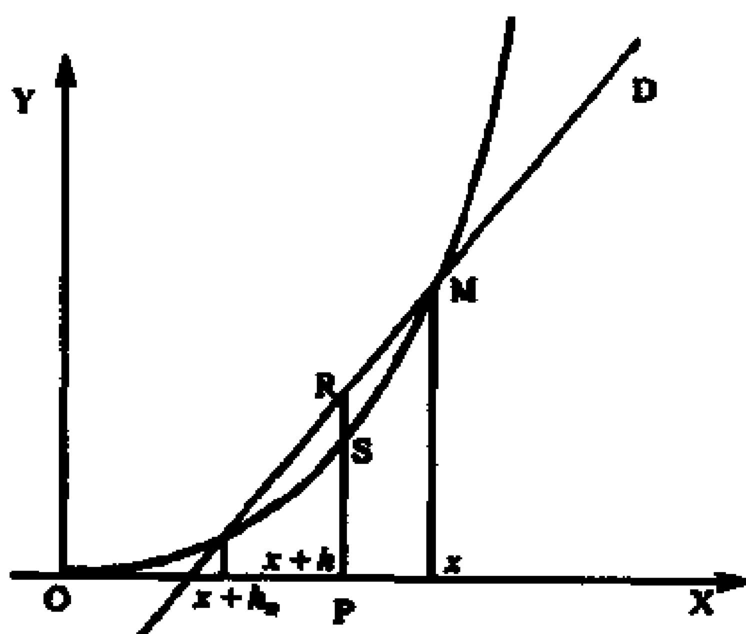


图 26

如果数 $\alpha = mx^{m-1} - a$ 不等于零，则最后式子中的括号与 α 正负号相同，只要

$$|h| < 1, |h| < |\alpha|/A,$$

而且不管 $h<0$ 或 $h>0$，情形都是如此；这样，随同 h 变得大于或小于零，$\overline{RS}$改变正负号. 换言之，D 在点 M 处穿过所给曲线. 因此，唯一的“切触”直线必是使 $a = mx^{m-1}$的直线 D_0. 一般地说该直线确实不穿过所给曲线，但 $m=3, x=0$ 的情形表明并非总是如此（图 27）. 然而，说方程为

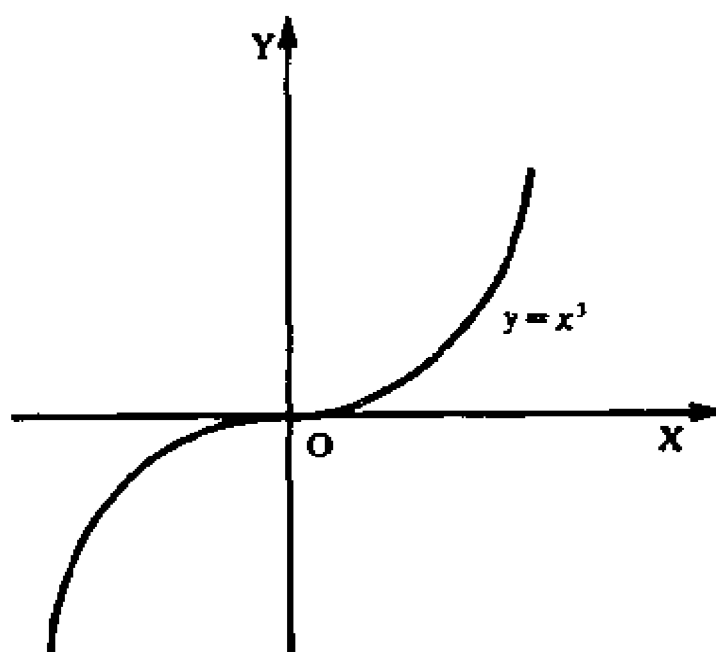


图 27

$$Y - x^m = mx^{m-1}(X - x)$$

[73] 的直线 D_0 在点 M 处与所给曲线相切还是方便的.

除非用一种 18 世纪中期之前在进行证明时并不清晰而且又用了 80 年之久才得到一般化的语言,无论是上述演算还是"穷竭法"论证,都不能提供令人满意的表述.这种语言就是极限语言.上述欧几里得的作图使得当 n 无限增大时内接于圆的正 2^n 边形"趋于"该圆"直观上显然成立".公元前 5 世纪,智者安梯丰走得如此之远,说对于充分大的 n,此内接正多边形变成圆本身,这使柏拉图学派的几何学家极为愤慨.后者手里有另一个例子即梅内克缪斯可能就已知道的双曲线的渐近线(图 19):点 $(x, 1/x)$ 到直线 OX 的距离小于 $1/n$,只要 $x > n$,但它决不会等于零①,这当然是只能对于柏拉图的"绝对图形"才能想象的

① 非数学家很难把握这个陈述.蒙田把它作为"学者"的怪癖来援引.这类概念上的困难可以帮助我们理解为什么达到极限一般概念要花那么长的时间.

事情.如今我们说实数列(h_n)以0为极限或趋于0,如果对于每个数$\alpha>0$,其下标n在某个数n_0以后的所有h_n都属于区间$-\alpha<x<\alpha$(n_0当然依赖于α).更一般地,我们说(h_n)趋于实数a或以a为极限,如果序列(h_n-a)趋于0;此时写

$$\lim_{n\to\infty} h_n = a.$$

用此种语言,区间套公理(§6)可表述为:如果

$$\lim_{k\to\infty}(c_k-b_k)=0,$$

则序列(b_k)与(c_k)具有相同极限,即含于所有区间中的唯一的实数Z.

回到费马确定曲线$y=x^m$的切线的演算,我们就能说点M处切线的"斜率"$a=mx^{m-1}$是连接点M与横坐标为$x+h_n$($h_n\neq0$)的点的"割线的斜率"

$$\frac{1}{h_n}((x+h_n)^m-x^m) \tag{5}$$

(它只当$h_n\neq0$时才有定义)对于每个趋于0的非零数序列(h_n)的极限.这种形式的定义可推广到任何方程为$y=f(x)$的曲线,条件是对于所有趋于0的非零数序列(h_n),序列

$$\frac{1}{h_n}(f(x+h_n)-f(x)) \tag{6}$$

[74]

存在极限.

在17世纪中,很快认识到费马的方法可应用到当时已知的许多其他曲线,例如$x^3y^2=1$,$y=\sin x$,$y=\tan x$,等等.对于当时通过伽利略和开普勒的工作已经开始出现的近代运动学和动力学,同样这种"趋于一个极限"的想法提供了"瞬时速度"——那时已经直觉地知道但尚未精确表述的概念——以及尤为重要的加速度概念的定义,它是牛顿力学的基石.

一定不能设想上面简短的概述反映了17世纪数学家们创造微积分的思维进展过程;这个过程是缓慢、曲折和混乱的.在表达式(5)中,当我们取$h_n=0$时,分子与分母两者都变为0,而

表示式 0/0 没有意义.数学家试图通过像莱布尼茨所说的"无穷小"[①]或像牛顿所说的"消逝量的最终值"来找到出路,但这只不过是用言词来掩盖观念的不精密.然而,到 17 世纪末,对于这类推理的逐渐熟练及其所取得的显著成果导致通过引进通用记号和算法对它加以整理编定,使之能容易地进行操作.当 x 变动时,表达式(6)的极限定义一个新的函数,牛顿把它称为函数 $y=f(x)$(他称之为"流")的"流数"并对这个新函数赋以记号 $\dot{y}$,至于莱布尼茨,则总是把它写为 $\mathrm{d}y/\mathrm{d}x$.后来(6)的极限被写为 $f'(x)$,并说这是函数 $y=f(x)$在点 x 处的导数.至于图 24 中的面积 S,从傅里叶时代起,就用函数 $y=f(x)$在 a 与 b 之间的积分这个术语和

$$\int_a^b f(t)\mathrm{d}t$$

这个记号,这是从莱布尼茨所用的名词直接继承下来的.此外,[75] 上述术语和记号也推广到在区间 I 上不一定取正值的函数,办法是把轴 OX 下方部分的面积取作负数.

由于导数与积分之间的联系这个关键发现,才使微积分的飞跃成为可能.这个联系是许多数学家各自独立地于 1660 年前后以多少有点混乱的方式发现的.当点 x 在区间 I 上变动时,式子

$$F(x)=\int_a^x f(t)\mathrm{d}t$$

定义了 x 的一个函数.假定有前面所作的假设,则函数 $y=F(x)$在每个点 x 处具有等于$f(x)$的导数(我们称 $F(x)$是$f(x)$

---

① 现代公理式方法能使我们严格地定义"非标准数"集 $\mathbf{R}^*$,它包含实数(也称为"标准数")集 $\mathbf{R}$,但在 $\mathbf{R}^*$ 中还有"无穷小数".通过使用非标准数,莱布尼茨的推理完全能加以验证.它组成称为"非标准分析"的分支,常能提供标准分析中定理的比较简单的证明,甚至还能导致新的性质的发现.

的原函数).这个事实易于证明.对 $h>0$,我们有(图 28)

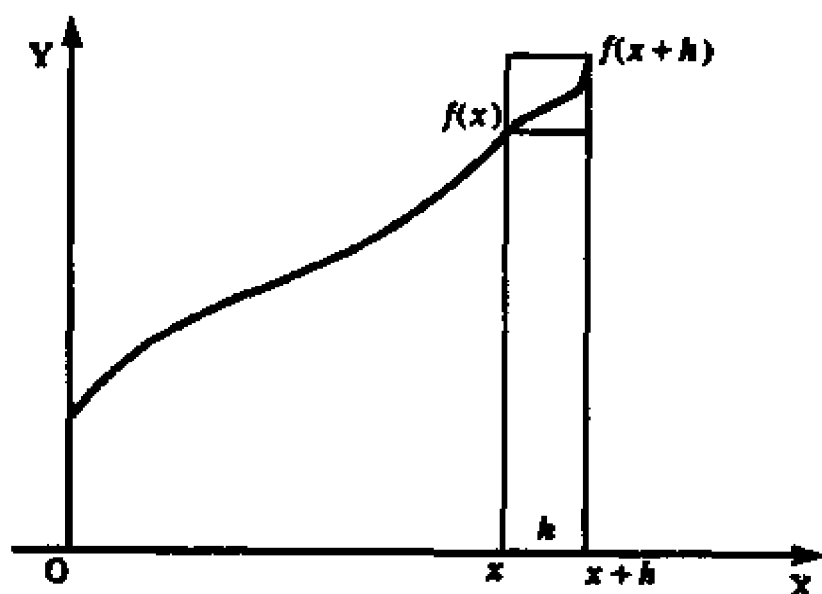


图 28

$$h\cdot f(x)\leqslant\int_{x}^{x+h}f(t)\mathrm{d}t\leqslant h\cdot f(x+h),$$

此式也可写为

$$0\leqslant\frac{1}{h}(F(x+h)-F(x))-f(x)\leqslant f(x+h)-f(x),$$

由假定 $f(x+h)-f(x)$ 随同 h 变为任意小①.对于 $h<0$,推理相同.

[76]

这个基本结论使得导数的性质移植为积分的性质以及反过来积分的性质移植为导数的性质成为可能.我们要在附录 5 中

---

① 一直到大约 1800 年,所有函数都不加指明地假定为(在现代意义下)连续的(附录 3).一旦得知函数 $f(x)$ 的一个原函数 $F(x)$,所有其余原函数就是 $F(x)+C$,其中 C 是任意常数.莱布尼茨用

$$\int f(x)\mathrm{d}x$$

记 f 的任意的原函数.

给出应用由此得到的算法的若干例子,它们使微积分成为异乎寻常灵活的工具.

最后,我们要注意,所有上面叙述的几何想法,像19世纪中叶之前微积分的所有其他几何应用一样,都基于关于曲线长度、曲面面积和立体体积的未予表述的假设之上.它们事实上是通过简化的图形由“几何直观”提示的.为把这些概念转化为柏拉图意[77]义下的“数学对象”,还有一些难题要在19和20世纪中来攻克.

附　　录

1. 欧几里得《几何原本》第Ⅴ卷中比的演算

欧几里得《几何原本》第Ⅴ卷讨论同类量之间的比,量所属的类则未予指明;在第Ⅵ和第Ⅹ到Ⅻ卷中,他把前面的理论应用于长度、平面图形面积和体积.对于同类的量(或如欧几里得所说“相互类同的”量),他使用第Ⅰ卷“共识”中引进的两种关系,用我们的记号就是 $A < B$ 和 $C = A + B$.在第Ⅴ卷定义5中,他加之还认可两个同类量满足阿基米德公理(本章§5).对于两个同类量 A,B,他引进一个新对象——两者之比,对此他未加定义(定义3是并不合用的“假定义”,像关于点和线的定义一样(本章§4)).为简短起见,我们把它写为 A/B,但欧几里得没有用记号而总是说“A 与 B 之比”.下一步,对于这些比,他用同类量之间的关系 $A = B$ 和 $A < B$ 定义比之间的关系 $A/B = C/D$(定义6)和 $A/B < C/D$(定义8).直到19世纪,大多数数学家觉得这些定义复杂得没有必要,甚至不可理解.我要在这里解释这种表面上的复杂性由于不可公度比的存在是必需的.它说明了古希腊人思想的深刻和精致.

对于整数之间的比,相等关系 $m/n = p/q$ 等价于 $mq = np$,不等关系 $m/n < p/q$ 等价于 $mq < np$.但关系 $m/n < p/q$ 也可用下面的方式表述:存在整数之比 r/s,使得

$$m/n < r/s < p/q \quad (1)$$

(如果已知 $m/n < p/q$ 而要求出这样的比,只要取 $r = m + p$, $s = n + q$即可).用另一方式也可说成:存在两个整数 r, s,使得 $rq < sp$ 但 $rn > sm$.

如果比 $A/B = m/n$ 与 $C/D = p/q$ 是可公度的,则基于前面所说,我们可把关系 $A/B < C/D$ 表述为:

存在两整数 r, s,使得 $rD < sC$ 但 $rB > sA$.

然而在这种表述中,比 A/B 与 C/D 为可公度这一事实是不相干的.这恰恰就是第Ⅴ卷中关于任何比(可公度或不可公度)的定义 8.

现在我们怎样来定义相等关系 $A/B = C/D$? 可以简单地说既非不等关系 $A/B < C/D$ 又非不等关系 $C/D < A/B$,由此就有第Ⅴ卷中关于 $A/B = C/D$ 的定义 6:

对所有使得 $rD < sC$(对应地,$rD = sC$,$rD > sC$)的整数 r, s,有 $rB < sA$(对应地,$rB = sA$,$rB > sA$).

于是,从欧几里得开始,当古希腊数学家要证明量之间的两个比的等式 $A/B = C/D$ 时,他们运用所谓的*反证法*,它由两个通过揭示矛盾的证明构成:首先证明存在两个整数 r, s 使得 $rD < sC$ 和 $rB > sA$ 是错的,然后证明调换 A/B 与 C/D 的地位时 [78] 出现同样的错误.正是用这种方法,例如,欧几里得运用本章 § 9 中说的"穷竭"过程证明两个圆面积之比是其半径平方之比(第Ⅻ卷命题 2)以及(用欧多克索斯的穷竭法)圆锥体积与同底同高圆柱体积之比等于 1/3(第Ⅻ卷命题 10).

2. 实数系的公理式理论

在第六章中我们会看到,为什么由笛卡儿开创的使实数等同于直线上的点(本章 § 4)最终引起被"严格性"盘据心头的 19 世纪数学家的疑虑,并会看到他们怎样喜欢实数的"算术式定义".另一方面,一旦希尔伯特在几何学中提供的公理式方法(本

章§4)得到真正理解,它可应用于数学的许多其他领域就变得很明显.正是按照这种方式,代数学、拓扑学和拓扑向量空间理论于1930年前后建立起来(见第五章§5,A).然而,只是到最近才认识到,同样的方法也能应用到实数系,而且只要不涉及无矛盾性问题,它就能为整个分析建立基础并能避免"算术式定义"(第六章附录2)的复杂性.

这个理论比欧几里得几何学的公理系统更简单.它只有一类"本原对象",称为"实数",而在实数之间,只有三种"本原关系":

1) 两个实数之间的关系 $x \leqslant y$(也写为 $y \geqslant x$)("序");

2) 三个实数之间的关系 $z = x + y$("加法");

3) 三个实数之间的关系 $z = xy$("乘法").

实数系理论需要17条公理:

R1) 存在两个不同的实数.

R2) 如果 x, y 是两个实数,则或者 $x \leqslant y$,或者 $y \leqslant x$.

R3) 如果同时有 $x \leqslant y$ 和 $y \leqslant x$,则 $x = y$;反之亦然.

R4) 如果 $x \leqslant y, y \leqslant z$,则 $x \leqslant z$.

R5) $(x + y) + z = x + (y + z)$.

R6) $y + x = x + y$.

R7) 存在实数 α,使对每个实数 x 有 $\alpha + x = x$.

R8) 对于每个实数 x,存在实数 x',使得 $x + x' = \alpha$.

R9) 如果 $x \leqslant y$,则对每个实数 z 有 $x + z \leqslant y + z$.

R10) $x(yz) = (xy)z$.

[79] R11) $yx = xy$.

R12) 存在实数 ϵ,使对每个实数 x 有 $\epsilon x = x$.

R13) 对于每个不等于 α 的实数 x,存在实数 x'',使得 $xx'' = \epsilon$.

R14) 如果 $\alpha \leqslant x, \alpha \leqslant y$,则 $\alpha \leqslant xy$.

R15) $x(y + z) = xy + xz$.

对于每个实数 x，我们按

$$1\cdot x=x,(n+1)\cdot x=n\cdot x+x$$

通过关于 n 的归纳定义数 $n\cdot x$.

R16) 如果 $\alpha\leqslant x$ 且 $\alpha\neq x$，则对每个实数 y，存在整数 n，使得 $y\leqslant n\cdot x$（"阿基米德公理"）.

R17) 如果 $(b_k)_{k\geqslant0}$，$(c_k)_{k\geqslant0}$ 是两个实数无穷序列，满足：对每个 $k\geqslant0$ 有

$$b_k\leqslant b_{k+1}\leqslant c_{k+1}\leqslant c_k,$$

则存在实数 x，使对每个 k 有

$$b_k\leqslant x\leqslant c_k$$

（"区间套公理"）.

作为公理式理论展开方式的例子，下面是由上述公理推出的第一批性质的证明：

1) 满足 R7) 的数 α 是唯一的. 因为如果 α' 是使得 $\alpha'+x=x$ 对每个 x 成立的另一实数，则由 R6) 得

$$\alpha=\alpha'+\alpha=\alpha+\alpha'=\alpha'.$$

以后我们把满足 R7) 的唯一的数记为 0.

2) 对于每个实数 x，满足 R8) 的实数 x' 是唯一的. 因为如果 $x+x'=0$，$x+y=0$，则由 R5) 和 R6)，有

$$\begin{aligned}x'&=x'+0=x'+(x+y)=(x'+x)+y\\&=(x+x')+y=0+y=y.\end{aligned}$$

以后我们把满足 R8) 的唯一的数 x' 记为 $-x$. 由定义得 $-(-x)=x$.

3) 关系 $x+z=y+z$ 蕴涵 $x=y$. 因为

$$\begin{aligned}x&=x+0=x+(z+(-z))=(x+z)+(-z)\\&=(y+z)+(-z)=y+(z+(-z))=y+0=y.\end{aligned}$$

4) 对每个实数 x 有 $0x=0$. 因为由 R15)，

$$0x+yx=(0+y)x=yx=0+yx,$$

由此由 3) 得 $0x=0$.

由于由 R1)得知存在 $x\neq0$,因此由 R12)和 R4)得 $\varepsilon\neq0$. 于是 R13)中条件 $x\neq0$ 是必要的.

5) 如果 $xy=0$,则两个数 x,y 中至少有一为 0(我们把这个事实说成不存在成为"零因子"的实数). 因为,如果 $x\neq0$,则由 R10),R11),R12)和 R13),有

$$x''(xy)=(x''x)y=\varepsilon y=y,$$

以及

$$x''(xy)=x''0=0x''=0.$$

我们把关系"$x\leqslant y$ 且 $x\neq y$"写为 $x<y$ 或 $y>x$. 于是,如果 $x<y$ 且 $y\leqslant z$,则有 $x<z$. 因由 R4)有 $x\leqslant z$,而如果 $x=z$,则会有 $y\leqslant x$ 从而由 R3)得 $x=y$,这与 $x\neq y$ 的假定矛盾. 同样可证,如果 $x\leqslant y$ 且 $y<z$,则 $x<z$. [80]

6) 如果 $x>0$,则 $-x<0$. 因为不可能有 $-x=0$,否则就会有 $x=-(-x)=0$. 也不可能有 $-x>0$,否则由 R9)就会有 $0=x+(-x)\geqslant x+0=x>0$,这是荒谬的.

7) 我们有 $(-x)y=-(xy)$,$(-x)(-y)=xy$(正负号规则). 因为由 2),4)和 R15),有

$$0=0y=(x+(-x))y=xy+(-x)y.$$

于是又有 $(-x)(-y)=-(x(-y))=-(-(xy))=xy$.

8) 对于每个 $x\neq0$,有 $x^2>0$. 因为若 $x^2=0$,则由 5),$x=0$. 如果 $x>0$,则由 R14)得 $x^2\geqslant0$,而 $x^2=0$ 又是不可能的. 如果 $x<0$,则 $-x>0$,从而由 7)得 $x^2=(-x)^2>0$.

9) 满足 R12)的数 ε 是唯一的.

10) 对每个 $x\neq0$,满足 R13)的数 x'' 是唯一的.

9)和 10)的证明过程与 1)和 2)的相同,只是要用乘法代替加法. 我们把满足 R12)的唯一的数记为 1,把满足 R13)的唯一的数记为 x^{-1} 或 $1/x$. 显然 $1=1^2>0$,因而 $-1<0$. 我们有 $(-1)x=-x$,因为

$$x+(-1)x=1x+(-1)x=(1+(-1))x=0x=0.$$

实数系的其他标准性质都能用同样方式证明.

对于每个自然数 n,把 n 等同于实数 $n\cdot 1$. 由实数 $0, n\cdot 1$ 和 $n\cdot(-1)=-(n\cdot 1)$(对于所有自然数 n)组成的集合 $\mathbf{Z}$ 称为整数集.

3. 多项式实根的逼近

本章 §6 中叙述的斯蒂文的方法基于下述引理(他并未证明):

如果多项式 $P(x)$ 对数 x_0 有 $P(x_0)>0$,则存在区间 $x_0-\alpha<x<x_0+\alpha$(其中 $\alpha>0$),使在该区间所有点处有 $P(x)>0$.

让我们假定此引理已得证,则连同 §6 中所用记号,我们能通过"反证法"(附录 1)证明 $P(Z)=0$. 先假定 $P(Z)>0$,则由引理,在区间 $Z-\alpha<x<Z+\alpha$ 内 $P(x)>0$ 也为真,但若取自然数 k 使得 $1/10^k<\alpha$,则由定义,$Z-\alpha<b_k\leqslant Z\leqslant c_k<Z+\alpha$,从而 $P(b_k)>0$,这与 b_k 的选法矛盾. 其次假定 $P(Z)<0$. 把引理应用于多项式 $-P(x)$ 可证必存在自然数 k 使得 $P(c_k)<0$,这又与 c_k 的定义矛盾. 这样我们确有 $P(Z)=0$.

上述引理可从下述陈述推出:

$$P(x_0+h)=P(x_0)+hP_1(x_0)+h^2P_2(x_0)+\cdots+h^nP_n(x_0),$$

其中 $P_1, P_2, \cdots, P_n$ 是多项式,此处不必给出其显式. [81]

如果 $|h|<1$,则我们有

$$|P_1(x_0)+hP_2(x_0)+\cdots+h^{n-1}P_n(x_0)|<M,$$

其中

$$M=1+|P_1(x_0)|+|P_2(x_0)|+\cdots+|P_n(x_0)|$$

是与 h 无关的不小于 1 的数. 这样我们有

$$\begin{aligned}P(x_0+h)&\geqslant P(x_0)-|h|\cdot|P_1(x_0)+hP_2(x_0)+\cdots+h^{n-1}P_n(x_0)|\\&\geqslant P(x_0)-M|h|.\end{aligned}$$

如果 $P(x_0)>0$,则由上式得到,对于满足 $-1<h<1$ 和

$$-\frac{P(x_0)}{2M} < h < \frac{P(x_0)}{2M}$$

的所有数 h,有 $P(x_0+h)>0$.

用类似论证可以证明,如果

$$P(t)=a_0t^n+a_1t^{n-1}+\cdots+a_n,$$

其中 $a_0>0$,则存在充分大的整数 $A>0$,使得 $P(A)>0$.事实上,对于 $t>0$,可写

$$P(t)=t^n\left(a_0+\frac{1}{t}\left(a_1+\frac{a_2}{t}+\cdots+\frac{a_n}{t^{n-1}}\right)\right).$$

如果 $t\geqslant 1$,我们有

$$a_0+\frac{1}{t}\left(a_1+\frac{a_2}{t}+\cdots+\frac{a_n}{t^{n-1}}\right)\geqslant a_0-\frac{1}{t}\left|a_1+\frac{a_2}{t}+\cdots+\frac{a_n}{t^{n-1}}\right|$$

$$\geqslant a_0-\frac{N}{t},$$

其中 $N=1+|a_1|+|a_2|+\cdots+|a_n|$.如果 N 是使得 $A>N/a_0$的整数,则就有

$$a_0+\frac{1}{A}\left(a_1+\frac{a_2}{A}+\cdots+\frac{a_n}{A^{n-1}}\right)\geqslant a_0-\frac{N}{A}>0,$$

从而作为 A^n 与一正数之积的 $P(A)$也是正数.

显然斯蒂文的方法可应用于任意的满足下述条件的函数 $f(t)$:$f(0)<0$,且对充分大自然数 $A>0$ 有 $f(A)>0$,只要假定上述关于多项式的引理对该函数也成立.对于我们称为连续的所有函数,情形正是如此;函数 $f(x)$连续是指,对它的定义区间的每个点 x 和每个数 $\alpha>0$,存在数 $\beta>0$(β 依赖于 x 和 α),使当 $-\beta<h<\beta$ 时有

$$|f(x+h)-f(x)|<\alpha.$$

事实上,如果 $f(x_0)>0$,只须在上面定义中取 $\alpha=f(x_0)/2$ 以保证当 $-\beta<h<\beta$ 时有

$$f(x_0+h)-f(x_0)\geqslant -\frac{1}{2}f(x_0),$$

从而就有 $f(x_0+h)\geq f(x_0)/2>0$. 可以证明连续性的一个等价表述是：对于每个趋于 0 的序列 (h_n) 有 $\lim\limits_{n\to\infty}(f(x+h_n)-f(x))=0$. 在 18 世纪末以前，数学家们都隐含地假定他们研究的所有函数都是连续的.

[82]

4. "穷竭法"论证

欧几里得描述的"穷竭法"(《几何原本》第Ⅻ卷，命题 2)，基于求出由圆的一条弦与弦的端点处的切线所构成的三角形 ACB(图 29)的面积的值. 此面积是三角形 ACH 面积的两倍，而后者等于

$$\frac{1}{2}AH\cdot CH=\frac{1}{2}AH^2\tan\alpha=\frac{1}{2}\sin^2\alpha\tan\alpha=\frac{1}{2}\frac{\sin^3\alpha}{\cos\alpha}.$$

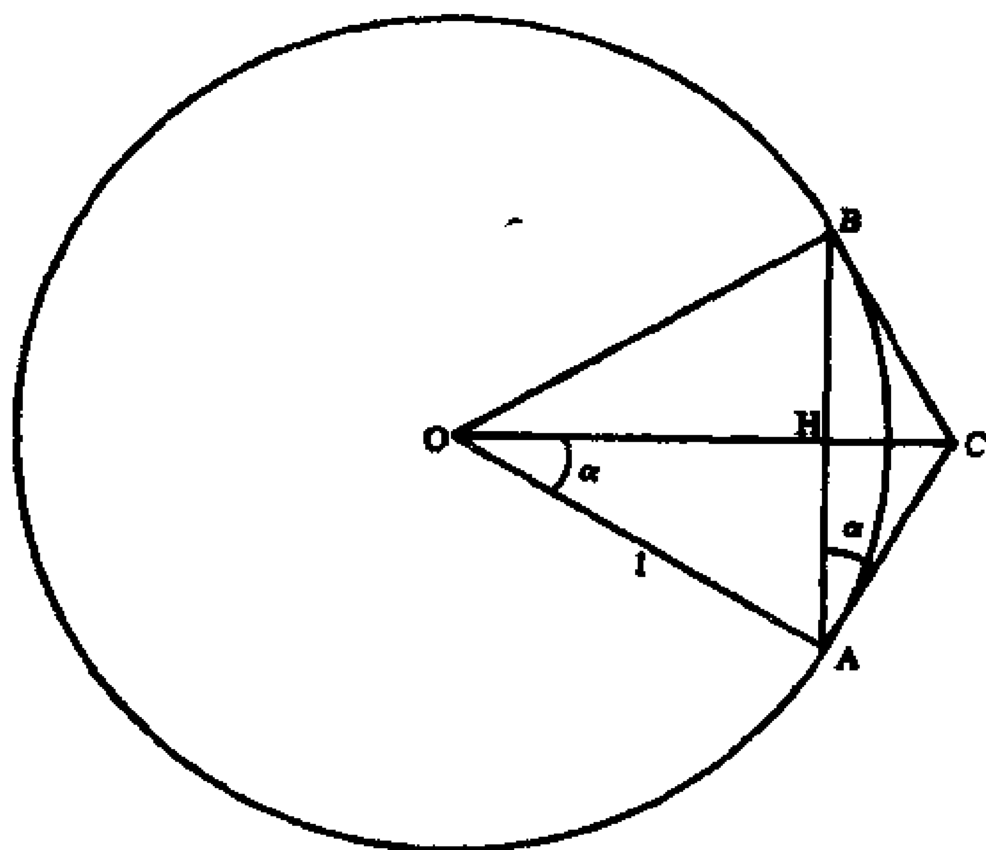


图 29

这样，欧几里得关于当圆弧 AB 二等分时相应三角形的面积减少一半以上的证明就归结为对于 $\alpha<\pi/4$ 的三角不等式

$$\frac{\sin^3\alpha}{\cos\alpha}<\frac{1}{4}\,\frac{\sin^3 2\alpha}{\cos 2\alpha}=2\,\frac{\sin^3\alpha\cos^3\alpha}{\cos 2\alpha}. \qquad (1)$$

它等价于 $\cos 2\alpha<2\cos^4\alpha$，最后又等价于

$$2\cos^4\alpha-2\cos^2\alpha+1>0,$$

而这个不等式由显见的不等式

$$2\left(x^2-\frac{1}{2}\right)^2+\frac{1}{2}>0$$

[83] 导出.

阿基米德用欧几里得的命题证明圆面积"等于一直角三角形的面积，其一条直角边与圆周同长，而另一直角边与该圆半径同长". 他就这样毫不犹豫地谈论曲线一段弧的长度(对此欧几里得学派是不予讨论的)，他而且意识到这就需要一些新的公设；基本的一条是说，包含于三角形 ABC 中的凸曲线弧 Γ(其中 B,C 是 Γ 的两端点，如图 30)的长度，等于介于 BC 与 $AB+AC$ 之间的一个线段的长度.

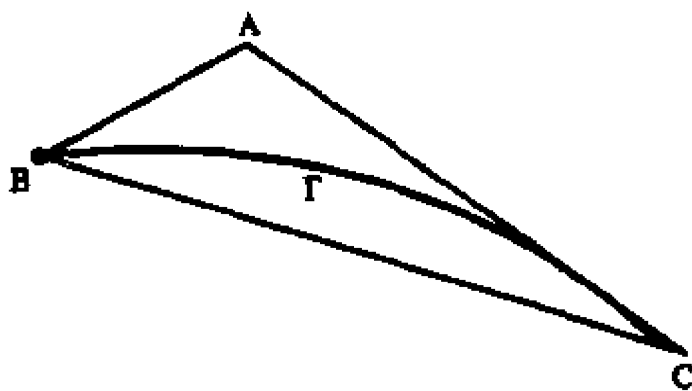


图 30

使用本章 § 9 中所用的记号，如果 l_n, L_n 分别是半径为 1 的圆的内接和外切正 2^n 边形的周长，则阿基米德就能写出不等式

$$l_n\leqslant 2\pi\leqslant L_n.$$

不过在关于圆周长度的短篇论著中,他只限于由圆的内接和外切正 96 边形周长给出的近似值.

人们会期望阿基米德像欧几里得那样,通过证明不等式

$$L_n - l_n \leqslant \frac{1}{2}(L_{n-1} - l_{n-1}) \tag{2}$$

给出 2π 作为周长 l_n 和 L_n 的"极限".

在图 29 中,我们有

$$(AC + BC) - AB = 2(\tan\alpha - \sin\alpha),$$

从而不等式(2)是对于 $\alpha < \pi/4$ 的三角不等式

$$\tan\alpha - \sin\alpha \leqslant \frac{1}{4}(\tan 2\alpha - \sin 2\alpha) \tag{3}$$

的推论,此不等式可写成

$$\frac{\sin\alpha(1-\cos\alpha)}{\cos\alpha} \leqslant \frac{1}{4}\cdot\frac{\sin 2\alpha(1-\cos 2\alpha)}{\cos 2\alpha} = \frac{\sin\alpha\cos\alpha(1-\cos^2\alpha)}{\cos 2\alpha}.$$

作一些初等变换,它就归结为

$$\cos 2\alpha \leqslant \cos^2\alpha(1+\cos\alpha),$$ [84]

最后归结为

$$\cos^3\alpha - \cos^2\alpha + 1 \geqslant 0.$$

当 $0 \leqslant x \leqslant 1$ 时,我们确有 $x^3 - x^2 + 1 \geqslant 0$.

然而,三角学演算是托勒密(公元 2 世纪)后才变得流行起来的,而不等式(3)的纯几何证明看来对于阿基米德又过于困难.

至于对自然数 m 计算原函数

$$\int_0^x t^m \mathrm{d}t,$$

正如阿基米德在 $m=2$ 的情形所看到的,它归结为求出和

$$\frac{x^{m+1}}{n^{m+1}}(1^m + 2^m + \cdots + (n-1)^m) \tag{4}$$

的值.

(4) 中括号内的数等于 $S_m(n)$, $S_m(t)$是有理系数 $m+1$

次多项式.此点可通过把该多项式写为

$$S_m(t)=a_0t^{m+1}+a_1t^m+\cdots+a_mt$$

并使

$$S_m(n+1)-S_m(n)=n^m$$

来证明.由此我们可确定系数 a_j(见第四章附录 4).特别是,我们得到 $a_0=1/(m+1)$.这样,(4)就是 $x^{m+1}/(m+1)$与另外的 m 个项之和,这些项当 n 无限递增时都趋于 0.费马的结论可由此推出.

5. 初等积分学的应用

为有益于对微积分基础有点熟悉的读者,我想给出几个例子,表明能从哈代判断为"单调枯燥,令人厌烦"(第二章 §1)的积分学的某些初等算法中导出什么.为简单起见,如同 18 世纪数学家那样,我们假定所考虑的函数都是连续的、分段单调的(即其定义区间可分为有限个子区间,在每个子区间上所给函数是递增或递减的,或者是一个常数).这些算法有 3 个:

1) 如果在$[a,b]$上 $f(x)\geqslant 0$,则有

$$\int_a^b f(x)\mathrm{d}x\geqslant 0,$$

而且仅当在$[a,b]$上 $f(x)\equiv 0$ 时才能有

$$\int_a^b f(x)\mathrm{d}x=0.$$

由此可推出,对于任意两个函数 f,g,如果在$[a,b]$上有 $|f(x)|\leqslant M$,则有

[85]

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leqslant M\int_a^b|g(x)|\mathrm{d}x. \tag{1}$$

2) 对$[a,b]$上的可微函数 u,v,有

$$\int_a^b u(x)v'(x)\mathrm{d}x=u(b)v(b)-u(a)v(a)-\int_a^b u'(x)v(x)\mathrm{d}x \tag{2}$$

("分部积分法").

3) 如果 ϕ 定义于 $[a,b]$ 上且是可微的，f 定义在包含 ϕ 所取的全部值的一个区间上，则有

$$\int_a^b f(\phi(t))\phi'(t)\mathrm{d}t = \int_{\phi(a)}^{\phi(b)} f(x)\mathrm{d}x \tag{3}$$

("变量置换法").

公式(2)，(3)不过是函数乘积和复合函数求导公式的另一写法.

18 世纪的数学家首先用这些公式得到"初等"函数的本身也是"初等"函数的原函数，但它们还有许多别的用途.（对于 17 和 18 世纪的数学家，"初等"函数概念并不是十分精确的. 通常这些函数由"常用"函数诸如 e^x，$\log x$，$\sin x$ 等通过反复进行代数运算和复合步骤得到. 例如，

$$(e^{x^2} + \sin x)^{1/3}$$

应看作"初等"函数. 在 19 世纪，刘维尔给出了更精确的定义并能证明某些"初等函数"如 e^x/x 等，其原函数不是"初等"函数.）

分析中最重要的公式之一——关于 $n+1$ 次可微函数的泰勒公式，来自反复运用公式(2)：

$$\int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\mathrm{d}t = -\frac{(x-a)^n}{n!} f^{(n)}(a) + \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\mathrm{d}t,$$

$$\cdots\cdots$$

$$\int_a^x (x-t)f''(t)\mathrm{d}t = -(x-a)f'(a) + \int_a^x f'(t)\mathrm{d}t,$$

$$\int_a^x f'(t)\mathrm{d}t = f(x) - f(a).$$

把这些等式相加给出

$$f(x) = f(a) + \frac{(x-a)}{1!} f'(a) + \cdots + \frac{(x-a)^n}{n!} f^{(n)}(a) +$$

$$\int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\mathrm{d}t.$$

下面是另外两个例子:

Ⅰ)在概率论中,对于位于“铃形曲线”下的面积,如果忽略变量的很大的值,就会产生误差.我们必须估计此误差,就是说必须计算积分

[86]

$$\int_x^\infty e^{-t^2}\mathrm{d}t.$$

这个 x 的函数不是“初等”函数,但进行分部积分,它可写为

$$\int_x^\infty e^{-t^2}\mathrm{d}t = \int_x^\infty \frac{1}{2t}\cdot 2te^{-t^2}\mathrm{d}t = \frac{e^{-x^2}}{2x} - \int_x^\infty \frac{e^{-t^2}}{2t^2}\mathrm{d}t,$$

又因对于 $x \leqslant t$ 有

$$1 \leqslant 1 + \frac{1}{2t^2} \leqslant 1 + \frac{1}{2x^2},$$

故得

$$\frac{e^{-x^2}}{2x\left(1+\frac{1}{2x^2}\right)} \leqslant \int_x^\infty e^{-t^2}\mathrm{d}t \leqslant \frac{e^{-x^2}}{2x}.$$

Ⅱ)在分析、数学物理和数论的很多问题中,常遇见形如

$$e^{if(t)} = \cos f(t) + i\sin f(t)$$

的函数,其中 f 是实值函数,常充当“相位函数”.点 $e^{if(t)}$ 位于半径为1的圆周上,按照 $f(t)$ 的快慢变化而在该圆周上快慢“旋转”.由于 $|e^{if(t)}| = 1$,所以由(1)给出的关于积分

$$\left|\int_a^b e^{if(t)}\mathrm{d}t\right|$$

的显然上界只是 $b-a$.事实上,这只当 f 为常数时达到.如果 $f(t)=ct$,则有

$$\int_a^b e^{ict}\mathrm{d}t = \frac{1}{ic}(e^{icb} - e^{ica}),$$

当 c 很小时,上式绝对值可以很大.这种情形下二阶导数 $f''(t)$

恒等于零.值得注意的是,当只知$|f''(t)|$的下界时,此积分的绝对值可以上方有界:如果在$[a,b]$上$f''(t)\geqslant\lambda>0$,则有

$$\left|\int_a^b e^{if(t)}\mathrm{d}t\right|\leqslant\frac{8}{\sqrt{\lambda}}$$

(范德科珀特不等式).

我们分几种情形来验证.

1) $b-a\leqslant 2/\sqrt{\lambda}$.此时由(1)有

$$\left|\int_a^b e^{if(t)}\mathrm{d}t\right|\leqslant\frac{2}{\sqrt{\lambda}}.$$

2) $b-a>2/\sqrt{\lambda}$,且在$[a,b]$上$f'(t)\geqslant 0$.此时对$a+\frac{2}{\sqrt{\lambda}}\leqslant t\leqslant b$有(因为$f'$递增)

$$f'(t)\geqslant\lambda\cdot\frac{2}{\sqrt{\lambda}}=2\sqrt{\lambda}.$$[1]

于是由(2)可写

$$\int_{a+2/\sqrt{\lambda}}^b e^{if(t)}\mathrm{d}t=\int_{a+2/\sqrt{\lambda}}^b\frac{1}{if'(t)}if'(t)e^{if(t)}\mathrm{d}t$$

$$=\frac{e^{if(b)}}{if'(b)}-\frac{e^{if(a+2/\sqrt{\lambda})}}{if'\left(a+\frac{2}{\sqrt{\lambda}}\right)}+\frac{1}{i}\int_{a+2/\sqrt{\lambda}}^b\frac{f''(t)}{(f'(t))^2}e^{if(t)}\mathrm{d}t,\qquad[87]$$

---

[1] 这是因为

$$f'\left(a+\frac{2}{\sqrt{\lambda}}\right)=f'(a)+\int_a^{a+2/\sqrt{\lambda}}f''(t)\mathrm{d}t$$

$$\geqslant\int_a^{a+2/\sqrt{\lambda}}f''(t)\mathrm{d}t\geqslant\lambda\cdot\frac{2}{\sqrt{\lambda}}=2\sqrt{\lambda},$$

从而当$t\geqslant a+\frac{2}{\sqrt{\lambda}}$时有

$$f'(t)\geqslant f'\left(a+\frac{2}{\sqrt{\lambda}}\right)\geqslant 2\sqrt{\lambda}.$$

——译注

而由(1)得

$$\left|\int_{a+2/\sqrt{\lambda}}^{b} \frac{f''(t)}{(f'(t))^2} e^{if(t)} \mathrm{d}t\right| \leqslant \int_{a+2/\sqrt{\lambda}}^{b} \frac{f''(t)}{(f'(t))^2} \mathrm{d}t$$

$$= \frac{1}{f'\left(a+\frac{2}{\sqrt{\lambda}}\right)} - \frac{1}{f'(b)}$$

$$\leqslant \frac{1}{\sqrt{\lambda}},$$

由此得到

$$\left|\int_a^b e^{if(t)} \mathrm{d}t\right| \leqslant \left|\int_a^{a+2/\sqrt{\lambda}} e^{if(t)} \mathrm{d}t\right| + \left|\int_{a+2/\sqrt{\lambda}}^{b} e^{if(t)} \mathrm{d}t\right| \leqslant \frac{4}{\sqrt{\lambda}}.$$

3) 一般情形.由于f''不改变正负号,所以f'只能在一个点$c(a \leqslant c \leqslant b)$处取零值.这样,由前述情形,

$$\left|\int_c^b e^{if(t)} \mathrm{d}t\right| \leqslant \frac{4}{\sqrt{\lambda}}.$$

另一方面,由(3)有

$$\int_a^c e^{if(t)} \mathrm{d}t = \int_{-c}^{-a} e^{ig(t)} \mathrm{d}t,$$

其中$g(t)=f(-t)$.由于$g''(t)=f''(-t)$,所以在区间$[-c, -a]$上有$g''(t)>\lambda$,且在此区间上$g'(t)=-f'(-t) \geqslant 0$.由前述情形,我们得到

$$\left|\int_a^c e^{if(t)} \mathrm{d}t\right| = \left|\int_{-c}^{-a} e^{ig(t)} \mathrm{d}t\right| \leqslant \frac{4}{\sqrt{\lambda}},$$

[88] 这就完成了(4)的证明.

第 四 章
经典数学中的某些问题

在第五章中，我们将试图表明，19 世纪的数学家们从经典的对象和方法出发，为使数学进一步发展，不得不发明新的数学对象和新的方法．

我想在本章和下章中简略地重温数学家们力图解决的某些问题(求解成功的程度则差异极大)的历史，这对理解这种必然性是有益的．在本章中，我们只讨论这样的问题，其最终解决不用到别的工具，只用到我们已讨论过的"经典"对象：数和几何图形．为使本书读者能够理解，我将把与称为数学分析即微积分概念的发展和多种应用这一分支(见第三章，§9)相联系的问题搁在一边；不过这些问题为数非常之多，而且是最引人入胜的壮观进展的主题．即使对某些可以完全用初等术语陈述的问题，富于想象力的数学家也通过引入来自数学分析的概念或技巧，成功地部分或全部解决了这些问题，而这些概念或技巧初看起来同要解决的问题毫无关系．除了间接提及这些方法外，我恐怕无法做得更多．这些方法使数学家们惊讶，使他们感到数学的深刻的、常常是神秘的统一性；在谈到这种统一性时，他们会毫不犹豫地使用"美"这个词．[89]

1．极难问题与不结果实的问题

A．完 满 数

某些迄今尚未解决的问题可追溯到古代．毕达哥拉斯学派

赋予自然数哲学而又神秘的概念,这导致他们观察到某些自然数具有这样的性质:这些数等于其所有因数之和[①].例如,

$$28 = 14 + 7 + 4 + 2 + 1.$$

他们把这种数称为完满数;问题在于求出所有完满数.古希腊数学取得非凡成就的证据之一,是欧几里得能够证明(《几何原本》,第Ⅸ卷,命题36),形如

$$N = 2^{p-1}(2^p - 1)(p \geq 2)$$

的数是完满的,如果假定 $2^p - 1$ 是素数.事实上,由于 $2^p - 1$ 是素数,N 的因数只有

$$1, 2, 2^2, \cdots, 2^{p-1}$$

以及

$$2^p - 1, 2(2^p - 1), \cdots, 2^{p-2}(2^p - 1).$$

但上述第一类因数之和是 $2^p - 1$,而第二类因数之和是 $(2^{p-1} - 1)\cdot(2^p - 1)$,因此 N 的全部因数之和是

$$2^p - 1 + (2^{p-1} - 1)(2^p - 1) = 2^{p-1}(2^p - 1) = N.$$

到了18世纪,欧拉证明这种数是仅有的偶完满数.这一结论立即引出两个问题:

1) 有没有奇完满数?

2) 对于何种自然数 $p(p \geq 2)$,$2^p - 1$ 是素数?

第一个问题至今仍无答案,尽管各种计算使人们倾向于设想其回答是否定的.至于第二个问题,很清楚自然数 p 本身必定应当是素数,因为如果 $p = mn$ 且 $m > 1, n > 1$,就会有

$$2^{mn} - 1 = (2^n - 1)(1 + 2^n + 2^{2n} + \cdots + 2^{(m-1)n}).$$

当 p 为素数时,形如 $2^p - 1$ 的数,称为梅森数.借助十分繁重的手工计算,验证出对于 $p \leq 127$,只有对应于 p 的值为

$$2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127$$

的梅森数是素数.

---

① 在这个陈述中,这些数自身不计入其因数之中.

电子计算机的降临使扩大这一清单成为可能，现在几乎每 [90] 年都能发现新的素梅森数．到 1985 年，最大的素梅森数是

$$2^{216091}-1,$$

它有 65050 位，是已知的最大素数①．但至今仍然不知道是否有无穷多个素梅森数．

B．费　马　数

看来费马是第一位对于数 2^k+1 提出类似问题的数学家．由于当 k 为奇数时多项式 T^k+1 被 $T+1$ 整除，因此数 2^k+1 中可能成为素数的只能是

$$2^{2^n}+1$$

这样的数(称为"费马数")．费马验证了

$$2^{2^0}+1=3,2^{2^1}+1=5,2^{2^2}+1=17,$$

$$2^{2^3}+1=257,2^{2^4}+1=65537$$

都是素数，从而他设想所有"费马数"$2^{2^n}+1$ 也都是素数．但欧拉指出

$$2^{2^5}+1=641\times6700417,$$

而到目前为止，还没有发现别的素费马数．

上述这些问题的特征不在于研究它们的过程中仍有未能弄清楚的关键之处，这其实是所有问题的特点；这些问题的特征在于，在研究中从未能再前进一步．当然，列举无数这种类似的极难问题是容易的．但是，为什么我们引述的这几个问题一直活跃在数学家的记忆中？"完满"数驱使我们回到算术的肇始年代，回到我们的先辈令人惊叹的年代；而梅森数和费马数则与著名数学家的姓氏联结在一起，而且最近它们以奇异的方式进入有

① 1998 年 1 月前发现的最大素梅森数是 $2^{3021377}-1$．——译注

限单群(第五章，§4,B)理论中，此外，高斯在他19岁那年作出的第一个伟大数学发现，就是具有 p 条边(p 是素数)的正多边形能用直尺和圆规作图当且仅当 p 是费马数.

再说，我们永远不应当绝望：50年前看来不可企及的问题后来却被解决了(见附录4).

C. 四色问题

这个问题大约提出于1850年，由于它可以非常简单地陈[91]述，还由于它极难解决，所以很快使非数学家着迷.一张地图画有 N 个国家，每个国家有一些邻国；相邻国家由边界(一条曲线弧)分隔(图31).四色问题就是要确定，四种颜色是否足以使 N 个国家都着上色并使得没有两个邻国具有相同的颜色.

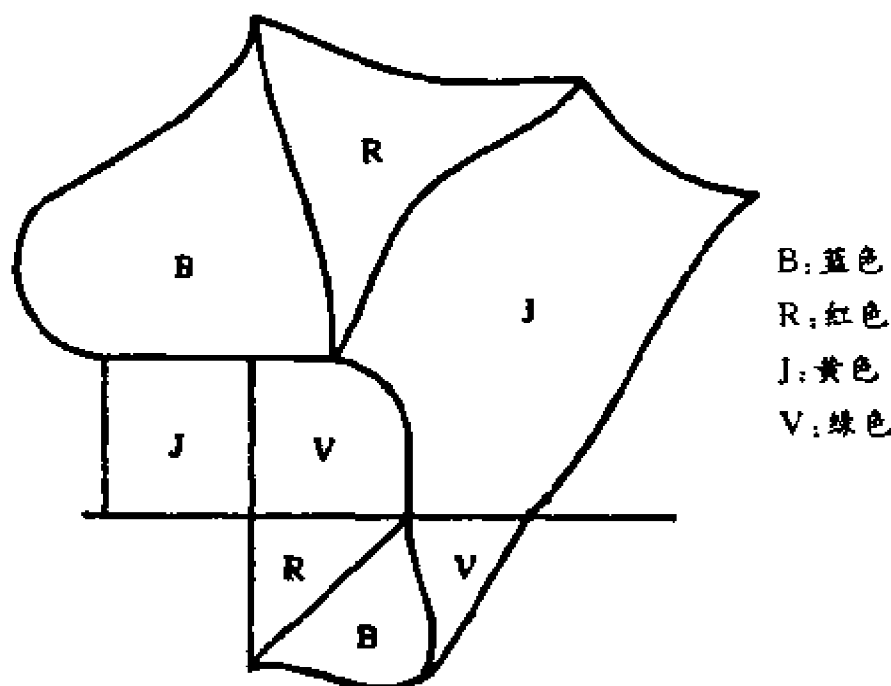


图 31

这个问题属于所谓组合论.在这个问题中，必须对 N 的每个值，鉴别 N 个国家与其他国家所有可能的位置关系.一个多

世纪里,在处理越来越大的数目 N 的过程中(例如 50 这样的数就已需要大量工作),出现了许多精巧独创的想法.值得注意的是,研究了一些典型位置,它们可用来把 N 个国家的问题简化为数目较小的 N' 个国家的问题.最后,阿佩尔和黑肯于 1977 年成功地证明,只须验证对于有限的 M 个适当确定的布局可以只用四种颜色.就能保证对于每个布局可以只用四种颜色.不过这个数目 M 超过 1800,从而"用手工"来验证所有这些情形的可能性是不能讨论的.为此必须求助威力巨大的电子计算机,经过一年的努力,这一定理终于借助计算机得证.看来已经做了几种独立的鉴定,因此我们可以认为四色问题已经解决,同时希望总有一天另外的方法会提供不那么笨重的证明.

同样,对于这么巨大的工作量,我们能说的只是看起来它与最终结论不相称.下面我们会看到(§2,B),在素数理论中还有一个未得证明的猜想,即黎曼假设.数论中有几十个结论依赖于 [92] 它:如果黎曼假设得证,那么这些结论就都得证;这一事实激发了数学家对它的兴趣.据我所知,四色问题没有任何这类现象.四色问题是一个孤立的结论,它未被用来证明"组合论"中别的定理,而且证明四色定理的方法也未被用来证明别的定理;当然我们不能谈及未来,但我们可以说,除非有新的结果出现,否则这是一个不结果实的问题.

D. 初等几何学中的问题

基于同样的理由,对于初等平面几何学中引起一代代数学教师和数学爱好者狂喜的绝大多数问题,也能下同样的结论.例如:

1) 卡斯蒂隆问题.给定一个圆 Γ 和任何不在该圆上的三个点 P,Q,R,要求作出一个三角形 ABC,使它内接于圆 Γ 且其边(可以延长)分别通过点 P,Q,R(图 32).

2) 马尔法蒂问题.要求作出三个圆,其中每一个与另外两个相切,同时与给定的三角形 ABC 的两条边相切(图 33).

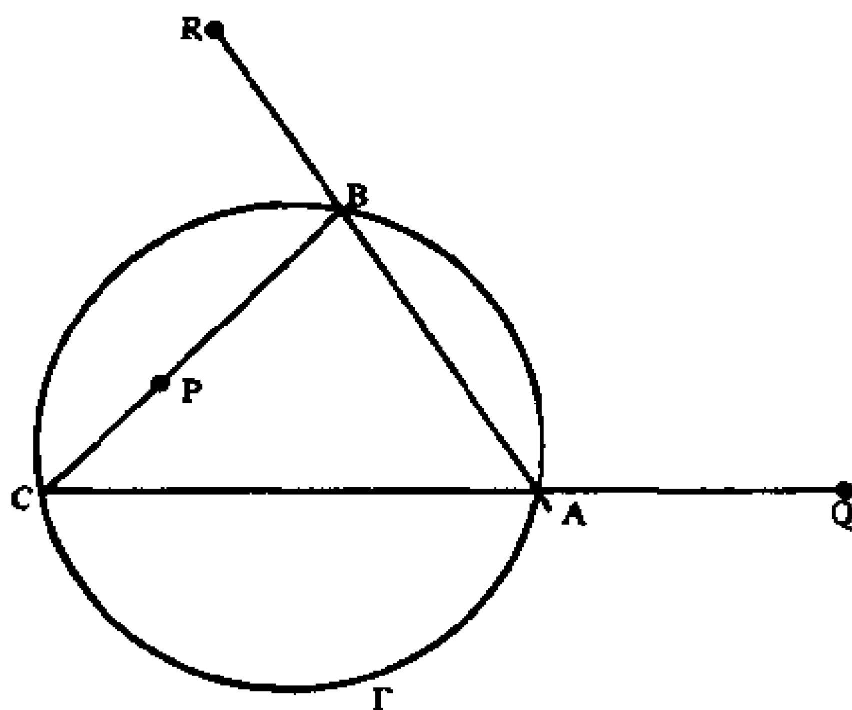


图 32

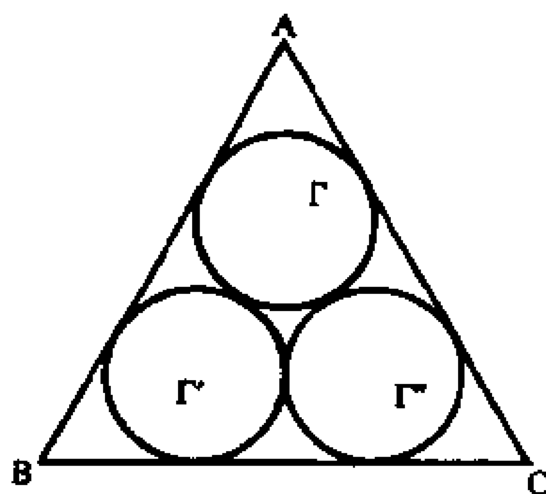


图 33

3) 圆规几何．这里的任务是求出只用圆规的作图方法，使之可以作出所有用直尺和圆规能作出的图形．

[93] 这些问题以及其他许多问题的解，散见于各种几何学论著中，它们或许有助于形成它已是一门完整的科学，其中再也没有

什么东西可发现的印象.矛盾的是,通常这些论著不提及"用直尺和圆规"不可能作出的几何问题,而它们却同理解数学的重大进步有着联系(第五章,§1).

2. 硕果累累的问题

A. 平 方 和

我们在第二章§5中提到了拉格朗日于1770年证明的下述定理:每个自然数是不超过4个自然数的平方之和.事实上,把一个自然数表示为一些自然数的平方之和这个问题在古典时代即已出现,无疑这与毕达哥拉斯定理有关,现在我们来解释这一点.

毕达哥拉斯定理导致寻求其边由自然数 a,b,c 度量的直角三角形,它等价于求出所有由3个自然数 a,b,c 构成的数组,满足

$$a^2=b^2+c^2.$$

令 $x=b/a,y=c/a$,它又等价于求出有理数 x,y,满足

$$x^2+y^2=1.$$

这是以 O 为圆心、半径为1的圆 C 的方程,且点 $A=(1,0)$ 位于此圆上.如果 $M(x,y)$ 是 C 上具有有理坐标的另一点,则连接点 M 与 A 的直线具有有理"斜率" t,其方程为

$$Y=t(X-1)$$

(图34).

如果我们写出此直线上的点 (x,y) 满足圆 C 的方程这个条件,即

$$x^2+t^2(x-1)^2=1,$$

则我们得到

$$(x-1)(x+1+t^2(x-1))=0;$$

由此得到点 M 的坐标为

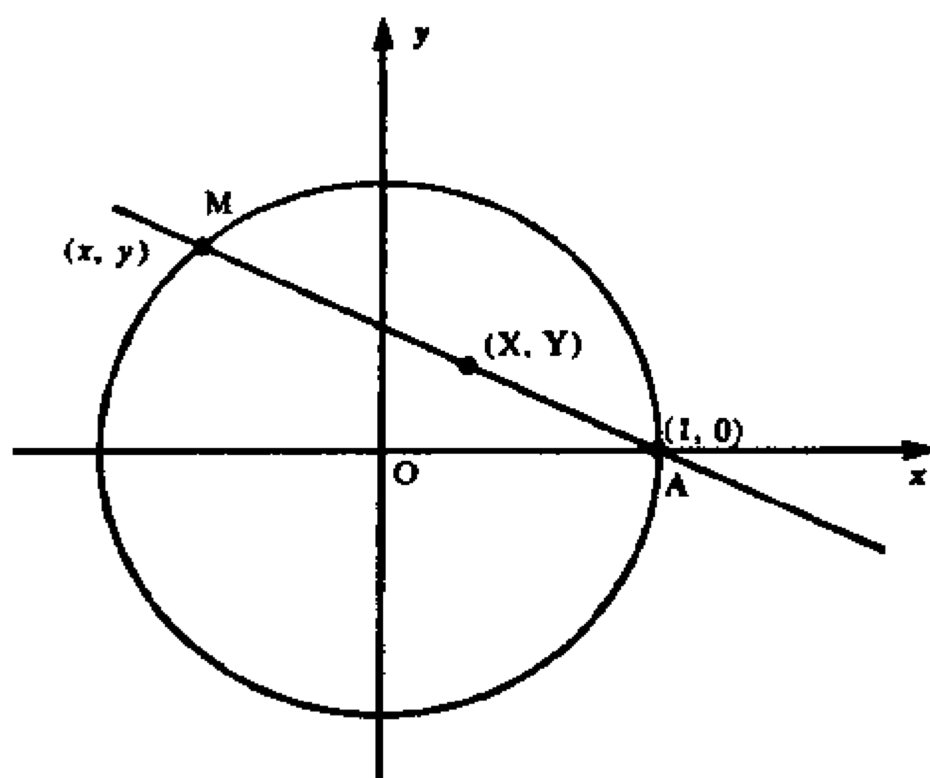


图 34

$$x=\frac{t^2-1}{1+t^2},\quad y=\frac{-2t}{1+t^2}.$$

如果取 t 为任一有理数,那么我们就得到 $x^2+y^2=1$ 的所有有理解.

[94] 丢番图即已知道形如 $4k+3$ 的数不可能是两个自然数平方之和,形如 $8k+7$ 的数不可能是 3 个自然数平方之和①,而且他可能已猜到拉格朗日定理.无论如何,在 17 世纪,巴谢·德·梅齐里阿发表过这个猜想,而费马则声称他证明了这个结果;他还断言每个形如 $4k+1$ 的素数只以一种方式成为两个平方之和,

① 每个数具有 $4k,4k+1,4k+2,4k+3$ 这四种形式之一.取这些数的平方,显见所得的数都具有 $4m$ 或 $4m+1$ 的形式,因此两个这样的数之和不可能具有 $4m+3$ 的形式.对于 3 个平方的情形,除了应考虑形如 $8k+h(0\leqslant h\leqslant 7)$的数的平方之外,推理过程相同.

但在他的论文中没有发现证明.上面第二个结论为欧拉所证明,他还成功地证明每个自然数是4个有理数的平方之和.另一方面,勒让德发表了这样的结果,即每个不是形如$4^m(8k+7)$的自然数是至多3个自然数的平方之和.他的证明并不完全令人信服,后来由高斯加以完善.

这里值得注意之处在于,上述定理一点也不使数学家避开这些显然已经解决的问题,而是激发大量新问题,对于这些新问题的研究构成数论中最广阔、最旺盛的分支之一;其中之一还同数学中别的领域,例如分析或几何相联系,而这种联系是最令人惊叹和最丰富多产的.可惜这些证明或者需要分析中的高等技巧,或者虽然初等(如费马和欧拉的证明)但却过于冗长和复杂,难以在这里复述(见附录3).因此我们只能限于陈述少数几个问题及其解. [95]

Ⅰ)除了两个或四个平方和,费马还研究过能写成a^2+Nb^2这种形式的自然数,其中a,b是自然数,在费马研究的情形中,$N=2$或$N=3$.欧拉证明并推广了费马对这个主题陈述但未给证明的一些定理.

正是拉格朗日从这些结果出发,开辟了一条通向惊人进展的新的途径.他以一般的方式考虑形如

$$ax^2+bxy+cy^2=n \tag{1}$$

的方程,其中a,b,c和n是整数(正或负的),而试图求出的解(x,y)中的x和y都是整数(正的或负的).(我们简单地把它们说成(1)的整数解.)拉格朗日的想法是在(1)中作"变量替换"

$$\begin{cases}x=\alpha X+\beta Y,\\ y=\gamma X+\delta Y,\end{cases} \tag{2}$$

其中$\alpha,\beta,\gamma,\delta$是整数.显然这一变量替换把方程(1)变换为类似的方程

$$AX^2+BXY+CY^2=n, \tag{3}$$

其中 A, B, C 都是(依赖于 $\alpha, \beta, \gamma, \delta$ 的)整数①. 如果我们有(3)的整数解(X, Y),那么显然公式(2)给出(1)的整数解(x, y). 然而,一般地说,不能由此得到(1)的全部整数解,因为由(2)解 X, Y 的两个线性方程构成的方程组,可导出

$$\begin{cases} X = \dfrac{\delta}{\alpha\delta - \beta\gamma} x - \dfrac{\beta}{\alpha\delta - \beta\gamma} y, \\ Y = \dfrac{\alpha}{\alpha\delta - \beta\gamma} y - \dfrac{\gamma}{\alpha\delta - \beta\gamma} x, \end{cases} \tag{4}$$

从而当 x, y 为整数时,X, Y 一般只是有理数. 除非

[96]

$$\alpha\delta - \beta\gamma = \pm 1, \tag{5}$$

这时 X, Y 必为整数. 于是拉格朗日限于考虑满足条件(5)的变量替换(2),此时他称多项式 $ax^2 + bxy + cy^2$ 与 $AX^2 + BXY + CY^2$ 是等价二次型. 后来高斯指出,保留这个概念是可取的;对于两个可由满足 $\alpha\delta - \beta\gamma = 1$ 的变量替换(2)互相变换的二次型,他称之为**真等价的**. 这样就能说相互之间真等价的二次型组成同一个类.

Ⅱ) 对于二元二次型 $ax^2 + bxy + cy^2$,数 $D = b^2 - 4ac$ 称为它的**判别式**. D 是一个整数. 同一类的所有二次型具有相同的判别式,这可由基于(2)的计算立即看出;但反之不然,具有给定判别式的二次型并不一定属于同一类. 然而,具有判别式 D 的类的数目 $h(D)$是有限的,它可明显地确定为 D 的一个函数;这是由狄利克雷作出的. 在 20 世纪,证明了确如高斯猜想的那样,

---

① 正是这种变量替换(其系数为任意)使得把表示圆锥曲线的方程(1)重写为

$$AX^2 + BY^2 = 1$$

这种形式成为可能. 但如果要求 $\alpha, \beta, \gamma, \delta$ 为整数且满足(5),那么并不总能选出 $\alpha, \beta, \gamma, \delta$,使得(3)中的"长方项"$BXY$ 为零. 例如,对于形式 $x^2 + xy + y^2$,我们有 $B = 2(\alpha\beta + \gamma\delta) + \alpha\delta + \beta\gamma$,而由 $\alpha\delta - \beta\gamma = \pm 1$,关系式 $B = 0$ 成为 $2(\alpha\beta + \beta\gamma + \gamma\delta) = \pm 1$,这当 $\alpha, \beta, \gamma, \delta$ 为整数时是不可能的.

当 D 趋向于 $-\infty$ 时，$h(D)$ 无限递增；另一方面，高斯还猜想当 D 趋向于 $+\infty$ 时，$h(D)$ 无穷次地取到值 1，但迄今仍不知道如何去证明或否定这个猜想.

仍不知道一般地说，为保证整数 n 使得方程(1)至少具有一个整数解(x, y)，何种判别准则是必要而又充分的. 拉格朗日和高斯只能给出一个涉及判别式 D 和整数 n 的判别准则，它保证对于*至少一个具有判别式 D 的二次型*，(1)至少有一个整数解. 不过如果相应于判别式 D 有若干个类，那就不可能指明对于哪个类(1)有整数解. 另一方面，拉格朗日表明，若(1)存在一个解，如何能从给定的一个解求出所有其他的解；如果 $D<0$，则有有限个解，但如果 $D>0$，则有无穷多个解.

上面谈到的所有结果都只涉及经典的对象和方法，但我们在第五章 §3, CⅡ)中会看到，它们如何成为现代数论的主要源泉之一，并如何促成新的数学对象——域、环和模的引进.

Ⅲ) 拉格朗日关于四平方之和的定理立即引起下面的新问题：方程

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = n \tag{6}$$

有多少整数解(x_1, x_2, x_3, x_4)?

此问题的答案由雅可比给出，它以出人意料的方式同 n 的算术性质相联系. 对每个指数 $k \geqslant 1$，我们以 $\sigma_k(n)$ 表示 n 的因数[①]的 *k 次幂之和*. 于是当 n 为奇数时(6)的解的个数为 [97] $8\sigma_1(n)$. 另一方面，如果 $n=2^r m$，其中 m 是奇数，$r \geqslant 1$，则(6)的解的个数为 $24\sigma_1(m)$.

自然这么好的开头不应就此打住，于是这个问题就推广为：对每个自然数 $m>4$，求出

---

① 此处我们认为一个数是它本身的一个因数. 这样前面所给的完满数定义相当于 $\sigma_1(n)=2n$.

$$x_1^2 + x_2^2 + \cdots + x_m^2 = n \tag{7}$$

的整数解$(x_1, x_2, \cdots, x_m)$的个数 $R_m(n)$.

对于这个问题没有适用于所有情况的完整结果.如果 $m = 8$,则我们有 $R_8(n) = 240\sigma_3(n)$,而如果 $m = 16$,则有 $R_{16}(n) = 480\sigma_7(n)$.一般地,如果 m 是 8 的倍数,则我们有 $R_m(n)$的下述"近似表示式":

$$\frac{m}{B_k}\sigma_{2k-1}(n), \tag{8}$$

其中 $k = m/4$, B_k 是第 k 个伯努利数(附录 4)."近似表示式"意味着

$$\varphi_m(n) = R_m(n) - \frac{m}{B_k}\sigma_{2k-1}(n)$$

与 $\sigma_{2k-1}(n)$相比非常之小;精确地说,显然我们有

$$\sigma_{2k-1}(n) > n^{2k-1},$$

而可以证明

$$|\varphi_m(n)| \leqslant C_m n^k,$$

其中 C_m 是一个不依赖于 n 的常数.

就其证明的方法来说,这些结论的证明仍然只需要 19 世纪的经典分析,但它们在群这个新概念(见第五章,§2,C)的源泉之中.

下一步是代替(7)考虑方程

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_m x^m + b_{12}x_1x_2 + b_{13}x_1x_3 + \cdots + b_{m-1,m}x_{m-1}x_m = n, \tag{9}$$

这是(1)的推广,其中系数 a_k 和 b_{ij}都是整数.关于它的理论遵循类似于方程(1)的理论的进程,但要复杂得多,并且需要更有威力的分析方法.

Ⅳ) 推广方程(6)的另一步是由华林跨出的,这件事甚至发生在拉格朗日的证明之前.华林问道:是否对每个 $k \geqslant 3$,存在只依赖于 k 的整数 $G(k)$,使对每个整数 n,存在方程

$$x_1^k + x_2^k + \cdots + x_{G(k)}^k = n \tag{10}$$

的整数解$(x_1, x_2, \cdots, x_{G(k)})$? 直到 1910 年,才由希尔伯特证明 [98] $G(k)$对每个 $k \geq 3$ 的存在性.从此以后,这个问题成为某些十分深刻的研究的主题,其间提出了(10)的解的个数的一些近似估计式.至于应当能推广二次型理论的"k 次型"的算术理论,尽管作了大量努力,至今仍处于萌芽阶段.对于由 n 个任意次方程构成的方程组,情形也是如此.

B. 素数的性质

据我所知,在公元前 5 世纪以前,除了希腊文明之外,没有任何别的文明有人想到过把一个自然数分解为它的素因数之积.如今我们把这种因数分解写为

$$n = p_1^{k_1} p_2^{k_2} \cdot \cdots \cdot p_r^{k_r}, \tag{11}$$

其中 $p_1, p_2, \cdots, p_r$ 是素数,k_i 是至少等于 1 的指数.由于没有适当的记号,它并未明显地出现于欧几里得的著作中.虽然如此,他却证明了下列三条性质(用现代语言表述):

a) 每个自然数或是素数,或被一素数整除(《几何原本》第Ⅶ卷,命题 31).

b) 如果 p 是素数,则 p^m 的因数只能是$p^r(r < m)$(《几何原本》第Ⅸ卷,命题 13).

c) 如果一个素数整除两个自然数之积 ab 且不整除 a,则它整除 b(《几何原本》第Ⅶ卷,命题 32).

由这些性质易于用归纳推理建立因数分解(11)的存在性和唯一性.

我们记得,在第二章 § 5 中,我们曾引述古希腊算术最优美的定理,即存在*无穷多个素数*.欧几里得给出的证明非常简单(参见哈代[7],第 28 页和本章附录 1),但我宁愿给出另一个属于欧拉的证明,因为它开辟了通向称为"解析数论"这一分支的道路;无论如何,它可以只用初等代数语汇来推述.

这里用的是"反证法",假定

$$p_1=2, p_2=3, p_3=5, \cdots, p_r$$

(按递增顺序排列)是仅有的素数,然后由此导出错误的结论.由(11),每个自然数 n 能且只能以一种方式分解为该式右端那样的乘积,这时应允许某些指数 k_i 等于 0(于是相应的因数 $p_i^{k_i}$ 代之以 1).取一个任意大的自然数 N,并考虑乘积

[99]

$$S_{N,r}=\left(1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^N}\right)\cdot\left(1+\frac{1}{3}+\frac{1}{3^2}+\cdots+\frac{1}{3^N}\right)$$
$$\cdots\cdot\left(1+\frac{1}{p_r}+\frac{1}{p_r^2}+\cdots+\frac{1}{p_r^N}\right). \qquad (12)$$

为计算这个乘积,只须在每个括号中取一项,计算所取的项之积,然后把所有这些部分积相加.这些部分积可写为 $1/n$,这里 n 取(11)所表示的形式,只是要限定对所有指数有 $0\leqslant k_i\leqslant N$.由因数分解的唯一性,所有这些部分积互不相同.关键之处在于,所有介于 1 与 2^N 之间的自然数 n 出现在一个部分积 $1/n$ 中(由上所述,它只出现一次).事实上,如果 $1\leqslant n\leqslant 2^N$,那么因数分解式(11)的指数 $k_2, k_3, \cdots, k_r$ 中没有一个能大于 $N-1$,因为否则 n 就会至少等于 3^N,与假定 $n\leqslant 2^N$ 矛盾.这样,$1/n$ 确实出现为 $S_{N,r}$表示式中的一个部分积.当然 $S_{N,r}$中有别的部分积,但我们所已证明的是

$$S_{N,r}\geqslant 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{2^N-1}+\frac{1}{2^N}. \qquad (13)$$

上式右端的和不易计算,但它能以下述方式代之以一个较小的数:把它的各项以截止到 1/2 的幂分成部分和组:

$$1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+\cdots+$$
$$\left(\frac{1}{2^{N-1}+1}+\frac{1}{2^{N-1}+2}+\cdots+\frac{1}{2^N}\right).$$

第一个括号中有 2 项,每项都至少等于 1/4,因而此括号中之和$\geqslant$1/2.第二个括号中有 4 项,每项都至少等于 1/8,因此它$\geqslant$

4/8 = 1/2. 依此类推,我们看出每个括号 $\geqslant 1/2$, 因为它有 2^{k-1} 项,每项都 $\geqslant 1/2^k$. 由于共有 $N-1$ 个括号,因此最后我们有

$$S_{N,r} \geqslant 1 + \frac{1}{2}N. \tag{14}$$

[100]

然而,通过给出几何数列之和的公式

$$1 + a + a^2 + \cdots + a^N = \frac{1 - a^{N+1}}{1 - a},$$

也能把 $S_{N,r}$ 表示为

$$S_{N,r} = \frac{\left(1 - \frac{1}{2^{N+1}}\right)\left(1 - \frac{1}{3^{N+1}}\right)\cdots\left(1 - \frac{1}{p_r^{N+1}}\right)}{\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\cdots\left(1 - \frac{1}{p_r}\right)}. \tag{15}$$

如果把分子中所有因数代之以 1,则我们就有不等式

$$S_{N,r} \leqslant \frac{1}{\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\cdots\left(1 - \frac{1}{p_r}\right)} = A_r, \tag{16}$$

而右边的项不再依赖于 N. 比较(14)和(16)得到

$$1 + \frac{1}{2}N \leqslant A_r,$$

从而 $N \leqslant 2(A_r - 1)$. 由于 N 可取得任意大,我们得到了矛盾的结果.

欧拉的想法是用幂 $(1/p_j)^m$ 代替(12)中素数的倒数 $1/p_j$, 这里指数 m 大于 1(但不必为自然数);这种想法是后来全部进展的根源. 可惜,除非使用分析中级数和无穷乘积概念,就不能描述这样得到的公式,而我们在这里不拟使用级数和无穷乘积(参见附录 2).

一旦建立了素数序列没有尽头这个事实,至少就有可能编制给出小于某个数的素数的表格. 相当早以前已有直至 $3\cdot 10^6$(三百万)的表格,而计算机则可以做得好得多. 已知的构造这种表格的最早方法是"爱拉托斯特尼筛法". 为找出 $\leqslant x$ 的所有素数,写下所有 $\leqslant x$ 的自然数构成的序列 $2, 3, 4, 5, \cdots, x$; 从 4 往下划掉 2 的

倍数,再从6往下划掉3的倍数,从10往下划掉5的倍数,依此继续下去.精确地说,在第 k 步后,没有划掉的 $k+1$ 个最小的数是素数,而如果 p_{k+1} 是其中最大者,则第 $k+1$ 步就是从 $2p_{k+1}$ 往下划掉 p_{k+1} 的倍数.我们会在最后一个素数 p_r 处停下,$p_r \leqslant \sqrt{x}$.事实上,如果自然数 m 满足 $\sqrt{x} < m \leqslant x$ 且未被划去,则它不可能是一个积 ab,其中 $a>1,b>1$,因为否则数 a 和 b 至少有一个会 $\leqslant \sqrt{x}$,因此 m 会被已经找到的素数之一整除从而应当会被划去.这[101]样,所有未被划去且 $>\sqrt{x}$ 的数是素数.

对于编制素数表,有着更有力的方法,但爱拉托斯特尼筛法或许促使欧拉考虑公式(16)中出现的数 A_r.事实上,对于素数 $p \leqslant \sqrt{x}$,有

$$\left(1-\frac{1}{p}\right)x$$

个数满足 $1 \leqslant m \leqslant x$ 且不是 p 的倍数.如果在"爱拉托斯特尼筛"中没有多次划去的数,那么近似地就会有

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\cdots\left(1-\frac{1}{p_r}\right)x \tag{17}$$

个小于 x 的素数,其中 p_r 是 $\leqslant \sqrt{x}$ 的最大的数.(17)中 x 的倍数正是数 A_r 的分母.这个数自然依赖于 x,而欧拉能够证明当 x 无限增大时此数趋向于零.然而,如果我们审察一下素数表,就会发现素数分布得很不规则.很多已知的素数 p 满足 $p+2$ 也是素数①(我们称它们构成一对"孪生素数"),人们甚至怀疑这样的素数对有无穷多个,尽管至今仍无方法证明此点.另一方面,在自然数序列中有你想多大就有多大的"洞",其中没有任何素数,例如序列

---

① 在1985年,我们知道有3424506个 $\leqslant 10^9$ 的数 p 满足 p 和 $p+2$ 都是素数*.

*)到1996年,已知共有182312485795对小于 1.37×10^{14} 的孪生素数;最大的孪生素数对是 $570918348\times10^{5120}\mp1$.——译注

$$n!+2, n!+3, \cdots, n!+n.$$

甚至欧拉都对这张表感到沮丧,觉得素数的分布是"人类心智永远不能看透之谜".然而,勒让德和高斯相互独立地在18世纪末叶提出了素数平均地遵从简单规律的想法.高斯考虑介于数 x 与数 $x+1000$ 之间的所有素数.如果 $N(x)$ 是这些素数的个数,那么他从素数表中观察到,当 x 很大时,比 $N(x)/1000$ 接近于 $1/\log x$.如果我们考虑以密度 $1/\log x$ 连续分布在一直线上的质量,那么介于2与 x 之间的质量会是

$$\mathrm{li}(x)=\int_2^x \frac{\mathrm{d}t}{\log t},$$

$\mathrm{li}(x)$ 称为积分对数.由此高斯设想,介于2与 x 之间的素数个数 $\pi(x)$ 必定能被数 $\mathrm{li}(x)$ 所"逼近".

[102] 通过猜想当 x 无限增大时比 $\pi(x)/\mathrm{li}(x)$ 趋向于1,上述想法就精确起来.这就是称为"素数定理"的命题.如果 $p_1, p_2, \cdots, p_n, \cdots$ 是素数构成的递增序列,那么可证明此定理等价于断言比 $p_n/n\log n$ 趋向于1.

对素数表的检验表明这些猜想是合乎情理的;例如,对于 $x=4\cdot10^{16}$,我们有

$$\pi(x)=1075292778753150, \mathrm{li}(x)-\pi(x)=5538861.$$

在整个19世纪中,有许多数学家致力于证明素数定理;但直到1896年,才由J·阿达玛和C·德·拉·瓦莱-普桑几乎同时予以证明.遗憾的是,所有证明方法都基于研究由黎曼引进的记作 $\zeta(s)$ 的函数,而这种研究需要分析中的高等技巧,以致我们不能加以叙述①.

① 1896年以后,许多数学家认为素数定理与黎曼 ζ 函数有不可分割的联系,因此不可能找到至多用一些初等微积分的"初等证明".然而,1949年,A·赛尔伯格和P·爱尔特希各自独立地给出了素数定理的初等证明,证明中除 e^x,$\log x$ 外不用任何"超越性"的东西,也不需要微积分.参见潘承洞、潘承彪,素数定理的初等证明,上海科学技术出版社,1988.——译注

数学家们并不满足于这个成就,他们还想了解差

$$\pi(x)-\mathrm{li}(x)$$

的性态.

黎曼关于函数 $\zeta(s)$ 有一个猜想.如果这个猜想为真,那么就会有下述结论:对每个指数 $\alpha>0$,

$$\frac{\pi(x)-\mathrm{li}(x)}{x^{\frac{1}{2}+\alpha}}$$

趋向于0.遗憾的是,尽管经历了130余年的奋斗,仍然无人能证明或否定黎曼假设.这个假设一直是数学中最重要的尚未解决的问题之一,因为它的解决会在数论的许多领域中带来巨大的进展.

在很长一段时间中人们相信会有 $\pi(x)<\mathrm{li}(x)$;素数表显示对于 $x\leqslant10^8$ 这一点成立.然而,李特尔伍德证明,存在无穷多个自然数 x,满足

$$\pi(x)-\mathrm{li}(x)>\frac{\sqrt{x}}{2\log x};$$

也有无穷多个自然数 x,满足

$$\pi(x)-\mathrm{li}(x)<-\frac{\sqrt{x}}{2\log x}.$$

迄今仍不知道使 $\pi(x)-\mathrm{li}(x)$ 改变符号的最小的数 x_0,但它必定非常之大.显然上述结果证实了素数的局部分布极不规则的印象.

1785年,由于想应用于二次型理论,勒让德需要改进欧几里得定理,即当 a 与 b 是两个没有公共素因数的(即 a 与 b 是互素的)自然数时,确定算术数列 $an+b$ 中是否存在无穷多个素数.对于某些特殊情形,通过以适当方式推广欧几里得的方法,可以容易地探明这一点.例如,对于算术序列 $4n+3$ 和 $6n+5$ 就能这样做(附录1).但是没有人能用初等方法成功地证明这个一般定理;此定理的证明是由狄里克雷于1837年给出的,证

[103]

明中运用了作为黎曼函数 $\zeta(s)$ 推广的函数.

我们以 $\pi(x;a,b)$ 表示算术序列 $an+b$ 中至多等于 x 的素数的个数.阿达玛和德·拉·瓦莱-普桑推广他们的方法,获得了 $\pi(x;a,b)$ 的一个估计式,并证明比

$$\frac{\pi(x;a,b)}{x/\log x}$$

当 x 无限增大时具有极限 $1/\varphi(a)$,其中 $\varphi(a)$ 是不超过 a 且与 a 互素的正整数个数.勒让德已猜想到这个公式.

C. 代数几何学的肇始

我们已看到(第三章,§8),在发明坐标方法之后,立即知道当 P 是二次多项式时由方程 $P(x,y)=0$ 给定的曲线是圆锥曲线.在18和19世纪,这导致按照多项式 P 的次数对方程 $P(x,y)=0$ 给定的曲线加以分类.例如,牛顿描述了三次曲线的各种形状.对于当 P 为多项式时由 $P(x,y,z)=0$ 给定的曲面,也以同样方式按 P 的次数加以分类.P 为二次时的曲面称为二次曲面,它们特别地包括球面和阿基米德研究过的某些旋转曲面.

在19世纪,许多数学家着迷于研究满足某些条件的圆锥曲线和二次曲面族,在这个领域中他们发现了许多优美的定理.低次曲线和曲面的几何研究也有一些可喜的成就.我们只引述一个例子:可以证明在三次曲面上或者有无穷多条直线——此时我们称它为直纹面,或者至多有27条直线,它们相互之间有着值得注意的几何关系.

这些研究是黎曼的工作开辟的现代代数几何学的前奏;在经历了一百多年的研究(在这种研究中分析、代数和拓扑都提供了重大贡献)之后,如今代数几何学已成为数学中扩展最为迅猛的领域.可惜这些主题具有高度技巧的本性,使我们不能在此予以讨论. [104]

附　录

1. 形如 $4k-1$ 或 $6k-1$ 的素数

不等于 2 的素数必定具有 $4k+1$ 或 $4k-1$ 两种形状之一，其中 $k \geq 1$. 稍微修改一下欧几里得的论证，可以证明有无穷多个形状为 $4k-1$ 的素数.

考虑形如 $4k-1$ 的素数构成的递增序列

$$p_1 = 3 < p_2 = 7 < \cdots < p_r, \tag{1}$$

假定它含有所有这类素数从而这类素数都 $\leq p_r$. 我们来证明存在形如 $4k-1$ 的不在这个序列中的素数，从而此素数 $> p_r$.

构造数

$$N = 4p_1p_2\cdots p_r - 1, \tag{2}$$

显然此数不能为序列(1)中任何数整除；另一方面，它不可能是形如 $4k+1$ 的素数之幂的乘积，因为每个乘积

$$(4a+1)(4b+1) = 4(4ab+a+b)+1$$

又有 $4c+1$ 的形状. 这样形如 $4k+1$ 的素数之积都不能等于由(2)定义的数 N，因此 N 必有一个形状为 $p=4k-1$ 的素因数，且 p 与序列(1)中每个数不同.

每个不等于 2 或 3 的素数必定具有 $6k+1$ 或 $6k-1$ 这两种形状之一. 在上述论证中以 6 代替 4，就可证明有无穷多个形状为 $6k-1$ 的素数.

2. 分解 $\zeta(s)$ 为欧拉积

设

$$p_1 < p_2 < \cdots < p_r < \cdots$$

是素数构成的(无穷)递增序列，并设 s 是一个大于 1 的指数. 关于素数的欧拉公式有如下述：

$$\left(1-\frac{1}{p_1^s}\right)^{-1}\left(1-\frac{1}{p_2^s}\right)^{-1}\cdots\left(1-\frac{1}{p_r^s}\right)^{-1}\cdots$$

$$=1+\frac{1}{2^s}+\frac{1}{3^s}+\cdots+\frac{1}{n^s}+\cdots, \tag{1}$$ [105]

每个自然数在右端只出现一次.为使此公式有意义,必须证明下列三点:

Ⅰ) 如果对每个整数 m 令

$$F(m)=1+\frac{1}{2^s}+\frac{1}{3^s}+\cdots+\frac{1}{m^s},$$

则序列$(F(m))$有极限 S.

Ⅱ) 如果令

$$G(r)=\left(1-\frac{1}{p_1^s}\right)^{-1}\left(1-\frac{1}{p_2^s}\right)^{-1}\cdots\left(1-\frac{1}{p_r^s}\right)^{-1},$$

则序列$(G(r))$有极限 P.

Ⅲ) $P=S$.

Ⅰ) 的证明.当 $m<m'$时,显然有 $F(m)<F(m')$.如同§2,B,我们对 $M<N$ 考虑差$F(2^N)-F(2^M)$,它可写为

$$\left(\frac{1}{(2^M+1)^s}+\cdots+\frac{1}{2^{(M+1)s}}\right)+\cdots+\left(\frac{1}{(2^{N-1}+1)^s}+\cdots+\frac{1}{2^{Ns}}\right).$$

每个括号

$$\frac{1}{(2^k+1)^s}+\frac{1}{(2^k+2)^s}+\cdots+\frac{1}{2^{(k+1)s}}$$

中有 2^k 个项,每项都小于 $1/2^{ks}$,因此该括号中各项之和小于 $1/(2^{s-1})^k$.由此可知 $F(2^N)-F(2^M)$ 小于首项为 $1/(2^{s-1})^M$,公比为 $1/2^{s-1}$的几何数列之和,从而

$$F(2^N)-F(2^M)\leqslant\frac{a}{(2^{s-1})^M}-\frac{a}{(2^{s-1})^N}, \tag{2}$$

其中

$$a=\left(1-\frac{1}{2^{s-1}}\right)^{-1}.$$

这样我们就能构造由区间

$$\left[F(2^M), F(2^M)+\frac{a}{(2^{s-1})^M}\right]$$

形成的区间套,其长度趋于零,且含有每个数 $F(m), m>2^M$. 由柯西原理,序列$(F(m))$存在极限 S,而 S 是含于所有这些区间中的唯一的数,并且对每个自然数 m 必有 $F(m)<S$.

Ⅱ)和Ⅲ)的证明. 当 $r<r'$ 时,有 $G(r)<G(r')$. 对每个自然数 N,存在自然数 r_N,使对每个自然数 $n\leqslant 2^N$, n 分解为素因 [106] 数之积的分解式中只含有素数 $p_1, p_2, \cdots, p_{r_N}$. §2,B)中的论证表明 $F(2^N)\leqslant G(r_N)$. 另一方面,对每个固定的自然数 r, $G(r)$是

$$\frac{\left(1-\frac{1}{p_1^{sM}}\right)\left(1-\frac{1}{p_2^{sM}}\right)\cdots\left(1-\frac{1}{p_r^{sM}}\right)}{\left(1-\frac{1}{p_1^s}\right)\left(1-\frac{1}{p_2^s}\right)\cdots\left(1-\frac{1}{p_r^s}\right)}$$

当 M 无限增大时的极限,而所写的这个数至多为 $F((p_1p_2\cdots p_r)^M)<S$. 这样,对每个 $r>r_N$,所有数 $G(r)$都在区间

$$[F(2^N), S]$$

中,而柯西原理表明序列$(G(r))$具有等于 S 的极限 P.

我们把(1)两边的共同值 S 记作 $\zeta(s)$.

3. 求 $ax^2+bxy+cy^2=n$ 的整数解的拉格朗日法

我们需要一些预备命题和注记.

Ⅰ)贝祖定理 如果 a, b 是两个自然数,则存在两个(正或负的)整数 m, n,使得

$$ma+nb=d, \tag{1}$$

其中 d 是 m 和 n 的最大公因数.

这个命题可由相继使用欧几里得除法得证. 如果 $a\geqslant b$,我们相继进行除法:

$$
\begin{cases}
a = bq + r_1, & r_1 < b, \\
b = r_1 q_1 + r_2, & r_2 < r_1, \\
r_1 = r_2 q_2 + r_3, & r_3 < r_2, \\
\qquad \cdots\cdots &
\end{cases}
\tag{2}
$$

由于余数 r_j 递减，我们必定会做到余数为零：

$$r_{k-1} = r_k q_k + r_{k+1}, r_{k+1} < r_k,$$

$$r_k = r_{k+1} q_{k+1}.$$

因为 a 和 b 的每个公因数必整除每个 r_j，所以有 $d = r_{k+1}$. 如果我们相继解(2)中每个方程，则会得到 $r_1 = a - bq, r_2 = b - r_1 q_1, \cdots$，由此看出所有 r_j 都有 $ua + vb$ 的形状，其中 u, v 是正或负整数.

Ⅱ）为解

$$ax^2 + bxy + cy^2 = n, \tag{3}$$

我们可以限于 a, b, c 的最大公因数等于 1 的情形. 因为如果 d 是 a, b, c 的公因数，那么 d 必定整除 n，而如果 $n = dn_1, a = d_1 a, b = db_1, c = dc_1$，那么方程(3)等价于 [107]

$$a_1 x^2 + b_1 xy + c_1 y^2 = n_1.$$

Ⅲ）我们可以限于求(3)的这样的解：x, y 的最大公因数等于 1（称为本原解）. 事实上，如果 $x = dx_1, y = dy_1$，则 d^2 必整除 n；令 $n = d^2 n_1$，我们就有

$$ax_1^2 + bx_1 y_1 + cy_1^2 = n_1.$$

Ⅳ）可以假定判别式 $D = b^2 - 4ac$ 不是整数平方. 因为否则方程 $at^2 + bt + c = 0$ 有两个有理根 $t_1 = v_1/u_1, t_2 = v_2/u_2$（写成既约分数形式，并取 u_1, u_2 为正数，而 v_1, v_2 或正或负），于是，由于

$$at^2 + bt + c = a(t - t_1)(t - t_2),$$

方程(3)就可写为

$$a(u_1 x - v_1 y)(u_2 x - v_2 y) = nu_1 u_2.$$

为解所得方程，我们以各种可能的方式把 $nu_1 u_2 / a$ 分解为

两个因数 r, s 之积,然后解方程组

$$u_1x - v_1y = r, u_2x - v_2y = s,$$

且限于求整数解(如果存在这种解的话).

V) 如果 b 是偶数,则判别式 D 是 4 的倍数;如果 b 是奇数,则 $D-1$ 是 4 的倍数. 当 b 为偶数时令 $\rho=0$,当 b 为奇数时令 $\rho=1$,于是 $(D-\rho)/4$ 总是整数.

VI) 我们把下述性质

$$a - b \text{ 是 } m \text{ 的倍数}$$

简记为"$a \equiv b(\mathrm{mod}\ m)$",并说成"$a$ 模 m 同余于 b".

拉格朗日定理于是就如下述:

为使具有判别式 D 的二次型 $ax^2+bxy+cy^2$ 使得方程(3)有一本原解,其必要充分条件是同余式

$$z^2 + \rho z - \frac{D-\rho}{4} \equiv 0(\mathrm{mod}\, n) \tag{4}$$

有一解 z,满足 $0 \leqslant z < n$.

必要性　设(α, γ)是(3)的一个本原解. 由贝祖定理,存在 β, δ,使得 $\alpha\delta - \beta\gamma = 1$. 如果作变量替换

$$\begin{cases} x = \alpha X + \beta Y, \\ y = \gamma X + \delta Y, \end{cases}$$

则得到方程

$$nX^2 + (2r + \rho)XY + sY^2 = n,$$

它有对应于(α, γ)的解$(1,0)$. 由于判别式不变,我们有

$$(2r + \rho)^2 - 4sn = D. \tag{5}$$

又由 $\rho^2 = \rho$,上式可写成

[108]

$$r^2 + \rho r - \frac{D-\rho}{4} = sn, \tag{6}$$

换言之,r 是同余式(4)的一个解,从而 r 除以 n 时所得余数 z 也是同余式(4)的一个解.

充分性　如果同余式(4)存在一个解 r,就是说存在一个

(正或负的)整数 s 满足(6),因而它满足(5),这意味着二次型

$$nx^2+(2r+\rho)xy+sy^2$$

以 D 为其判别式,且当 $x=1,y=0$ 时它取值 n.

让我们取方程

$$x^2+y^2=n \tag{7}$$

作为例子.

可以证明,只有一个二次型类以 $D=-4$ 为其判别式;这样拉格朗日定理表明方程(7)具有自然数解的必要充分条件是同余式

$$z^2\equiv-1(\bmod n) \tag{8}$$

有解.

假定 n 是一奇素数.我们应用费马的一条著名定理(我们会在第五章附录3中再次遇见它),即对每个整数 $a\neq0$,有

$$a^{n-1}\equiv1(\bmod n). \tag{9}$$

如果同余式

$$z^2\equiv a(\bmod n)$$

有一解,则必有

$$a^{(n-1)/2}\equiv z^{2(n-1)/2}\equiv1(\bmod n),$$

可以看出(见前述)这个条件是充分的.由于$(-1)^{(n-1)/2}$当 $n=4k+3$ 时同余于 -1 而当 $n=4k+1$ 时同余于1,我们得知每个形如 $4k+1$ 的奇素数是两个自然数平方之和,这就是由费马陈述并由欧拉证明的定理.

4. 伯努利数与 ζ 函数

这个附录是为在一定程度上熟悉通常幂级数展开式的读者写的.

我们在第三章§9及其附录4中看到,费马在最早计算原函数时,不得不追随阿基米德,对任意的自然数 $m\geqslant1$ 计算和式

$$S_m(n)=1^m+2^m+\cdots+n^m. \tag{1}$$

为此只须求出多项式 $S_m(t)$,满足方程

[109]

$$S_m(t+1)-S_m(t)=t^m. \tag{2}$$

从二项式公式可以直接验证,如果 $S_m(t)$是 $m+1$ 次多项式,则除常数项外,$S_m(t)$的系数完全由方程(2)确定.常数项由条件 $S_m(1)=1$ 给出,此条件则来自(1).

雅格布·伯努利给出了多项式 $S_m(t)$的优美描述,它可引述如下.对方程(2)稍加修改,对每个 $n\geqslant 1$ 我们考虑 n 次多项式 $\varphi_n(x)$,满足方程

$$\varphi_n(x+1)-\varphi_n(x)=nx^{n-1}. \tag{3}$$

我们考虑序列($\varphi_n(x)$)的"生成函数"

$$F(z,x)=\sum_{n=1}^{\infty}\frac{\varphi_n(x)}{n!}z^{n-1}, \tag{4}$$

这里的级数必须对 0 的一个邻域中的 z 和每个 x 收敛.由(3)得知必有

$$F(z,x+1)-F(z,x)$$
$$=\sum_{n=1}^{\infty}\frac{x^{n-1}z^{n-1}}{(n-1)!}=e^{zx}. \tag{5}$$

如果我们求(5)的形如

$$F(z,x)=e^{zx}f(z)+g(z)$$

的解,那么就会得到 $f(z)=1/(e^z-1)$. $g(z)$可取任意的函数.但是 $f(z)$在 $z=0$ 处没有定义,而通过令 $g(z)=-1/z$ 可以消除这一奇异性.这样就能通过关系式

$$\frac{e^{zx}}{e^z-1}=\frac{1}{z}+\sum_{n=1}^{\infty}\frac{\varphi_n(x)}{n!}z^{n-1} \tag{6}$$

定义伯努利多项式 $\varphi_n(x)$.

对于 $x=0$,这导致级数展开式

$$\frac{1}{e^z-1}=\frac{1}{z(1-u(z))}=\frac{1}{z}(1+u(z)+(u(z))^2+\cdots$$

$$+(u(z))^n+\cdots), \tag{7}$$

其中

$$u(z)=\frac{1+z-e^z}{z}=-\frac{z}{2!}-\frac{z^2}{3!}-\cdots-\frac{z^n}{(n+1)!}-\cdots. \tag{8}$$

由于

$$\frac{1}{e^z-1}=-\frac{1}{2}+\frac{e^z+1}{2(e^z-1)},$$

且

$$\frac{e^{-z}+1}{e^{-z}-1}=-\frac{e^z+1}{e^z-1},$$

所以

$$\frac{1}{e^z-1}+\frac{1}{2}$$

的级数展开式只含有奇次幂项.如果我们写

$$\frac{1}{e^z-1}=\frac{1}{z}-\frac{1}{2}+\sum_{n=1}^{\infty}(-1)^{n-1}\frac{B_n}{(2n)!}z^{2n-1}, \tag{9}$$

那么由(7)和(8)得知 B_n 都是有理数.计算给出前几个 B_n 的值如下: [110]

$$B_1=\frac{1}{6}, B_2=\frac{1}{30}, B_3=\frac{1}{42}, B_4=\frac{1}{30}, B_5=\frac{5}{66}, B_6=\frac{691}{2730},$$
$$B_7=\frac{7}{6}, B_8=\frac{3617}{510}, B_9=\frac{43867}{798}, B_{10}=\frac{174611}{330}, B_{11}=\frac{854513}{138},$$
$$B_{12}=\frac{236364091}{2730}, B_{13}=\frac{8553103}{6}, B_{14}=\frac{23749461029}{870}.$$

这些数称为伯努利数,它们在数学的不同领域诸如分析,数论(尤其在费马方程 $x^n+y^n=z^n$ 的研究中)和微分拓扑中有着重要的甚至是神秘的作用.如果把(9)与 e^{zx} 的幂级数展开式相乘,就能得到伯努利多项式的表示式

$$\varphi_n(x)=x^n-\frac{n}{2}x^{n-1}+\binom{n}{2}B_1x^{n-2}-\binom{n}{4}B_2x^{n-4}$$
$$+\binom{n}{6}B_3x^{n-6}-\cdots, \tag{10}$$

其中右边的和到使得指数非负的最后一项为止.我们有公式

$$\varphi_n(1-x)=(-1)^n\varphi_n(x), \tag{11}$$

$$\varphi_{2k+1}(0)=0, \varphi_{2k}(0)=(-1)^{k+1}B_k, \tag{12}$$

$$\varphi'_n(x)=n\varphi_{n-1}(x), n\geqslant 2. \tag{13}$$

它们可由(6)和方程

$$\frac{e^{z(1-x)}}{e^z-1}=-\frac{e^{-zx}}{e^{-z}-1}$$

直接推出.

在第五章附录5中,我们将证明,对于 $0\leqslant x\leqslant 1$,伯努利多项式有下列值得注意的三角级数展开式:对于 $k\geqslant 1$,

$$\varphi_{2k}(x)=(-1)^{k+1}2(2k)!\sum_{n=1}^{\infty}\frac{\cos 2n\pi x}{(2n\pi)^{2k}}, \tag{14}$$

$$\varphi_{2k+1}(x)=(-1)^{k+1}2(2k+1)!\sum_{n=1}^{\infty}\frac{\sin 2n\pi x}{(2n\pi)^{2k+1}}. \tag{15}$$

如果在(14)中令 $x=0$,我们就求得由欧拉获得的关于 $\zeta(2k)$ 之值的公式:

$$\zeta(2k)=1+\frac{1}{2^{2k}}+\frac{1}{3^{2k}}+\cdots+\frac{1}{n^{2k}}+\cdots=\frac{2^{2k-1}}{(2k)!}B_k\pi^{2k}, \tag{16}$$

这一公式曾使欧拉的同时代人大加赞许.这样,对于较小的 k 值,我们有

$$\zeta(2)=\frac{\pi^2}{2\cdot 3}, \zeta(4)=\frac{\pi^4}{2\cdot 3^2\cdot 5}, \zeta(6)=\frac{\pi^6}{3^3\cdot 5\cdot 7}, \zeta(8)=\frac{\pi^8}{2\cdot 3^3\cdot 5^2\cdot 7}.$$

老早就已知道 π 及其所有的幂都是无理数,因此对 $\zeta(2k)$ 同样的结论成立.由于对 $\zeta(2k+1)$ 没有任何类似于(16)的表示式,因此从欧拉时代以来就提出了 $\zeta(2k+1)$ 是否也是无理数的问题,而欧拉本人曾试图证明情形确是如此.二百年中在解决这[111]个问题上没有丝毫进展,因而这个问题显得是极难处理的.然而,法国数学家R·阿佩里于1977年想出了一个十分精巧的方法,它虽然只用到欧拉时代就已知道的结果,却能证明 $\zeta(3)$ 是无理数.遗憾的是,尽管进行了许多研究,还没有人能成功地找

到可以用于 $\zeta(5),\zeta(7)$等的类似方法.

回到费马原先的问题,从(3)我们求得

$$\varphi_{m+1}(1)-\varphi_{m+1}(0)=0,$$
$$\varphi_{m+1}(2)-\varphi_{m+1}(1)=(m+1)\cdot 1^m,$$
$$\varphi_{m+1}(3)-\varphi_{m+1}(2)=(m+1)\cdot 2^m,$$
$$\cdots\cdots$$
$$\varphi_{m+1}(n+1)-\varphi_{m+1}(n)=(m+1)\cdot n^m,$$

把这些等式的两边分别相加,我们得到

$$S_m(n)=\frac{1}{m+1}(\varphi_{m+1}(n+1)-\varphi_{m+1}(0)),$$

从而公式(10)给出了 $S_m(n)$通过伯努利数的明显表示式. [112]

第 五 章
新的对象和新的方法

18世纪是由于17世纪引进的数学技巧尤其是分析及其各种应用的精深发展而使人眼花缭乱的时代；对于分析对其他数学分支诸如几何学和概率论的应用以及对力学和天文学的应用(在预言自然现象方面，它取得了著名成就①)，情形也是如此.然而，奇怪的是，那个世纪数学似乎以陷于死胡同结束.那个世纪中叶的伟大数学家纷纷谢世，丹尼尔·伯努利殁于1782年，欧拉和达朗贝尔殁于1783年.拉格朗日还没有到50岁就相信纯粹数学进步的年代已经过去，1785年后他和蒙日都把注意力转向物理学和化学，而拉普拉斯则只关注力学和概率论.接着来到的是革命年代，它使科学工作者服务祖国，参与战争.其结果是1786至1796这十年内法国没有出现任何重要的数学新成果.像这样与社会动荡相联系的数学贫瘠年代，在当代也出现于几个国家中；不过在那个年代，除了法国之外，没有任何国家拥有足以与我们方才提到过的人相匹敌的卓越数学家，这就意味着当时数学的不毛状态是全球性的.

因此，1796年以高斯为标志所开创的是数学的真正复兴.高斯只通过研读前辈的著作来训练自己，他在15年内革新了全部数学.从19世纪早期起他不再孤独.法国大革命最有价值的

---

① 陀螺是一经典例证，不用刚体动力学方程就不能解释其自相矛盾的运动.

改革之一是在科学领域创建了真正的高等教育，它以优秀的教师为支柱，并向公众开放[1].综合工科学校——虽然其任务主要 [113] 是培训军事人才和工程师——在四分之三个世纪中始终是第一流数学家、物理学家和化学家的摇篮.除高斯外，直到 1825 年，由综合工科学校培养出来的数学家是无人可与之相比的.

正是 18 世纪甚至更早形成的观念构成了这一数学复兴的基础.但几乎同时，数学的风格和内容却改变了，这不仅是由于由高斯、波尔查诺、柯西和阿贝尔倡导的著名的“返回严格性”运动(见第六章，§2，A)，而且是由于引进了新的数学对象，它们与经典的对象大不相同，因为它们不再能用人们感官所能感知的“形象”来表示.

在整个 19 世纪，新观念的繁殖在数学的所有传统领域——算术、代数、几何、分析——中势不可当地持续进行.现在我们把 19 世纪看作一个过渡阶段，它构成“经典”数学与现代数学之间的桥梁.这是令人惊异的丰产年代：一方面发现了许多概念，它们成为数学的崭新领域——群论、拓扑学、函数空间等等——的基础，这些领域在当代得到了可与上述传统领域媲美的扩充；另一方面深入挖掘了许多旧概念，使之有可能理解有关的各种公理的真正意义.

一个一般观念逐渐显现，这就是作为一种数学理论基础的结构这个一般观念，它要到 20 世纪才能精确化.它起源于这样的观察：在一种数学理论中，起着根本作用的是所涉及的数学对象之间的关系，而不是这些对象的本性；因此在两个很不相同的理论中，却可能用相同的方式表述两者各自的关系.这些关系及其推论的系统形成一个隐藏在这两种理论深处的同一结构.

通过若干可以理解的例子，读者在本章中将会看到，一些巨

[1] 1789 年前的王朝只有专门培养未来官吏的学校才开设教一点微积分的数学课程，而普通百姓很难进入这样的学校.

大的结构如何通过 19 世纪的数学进程涌现出来，这些结构已是当代数学的基础. 这里必须强调，几乎在所有情形，这些结构的出现是基于解决成功地攻克从经典数学继承下来的问题的需[114] 要，而不是归于创造没有鲜明目标的新抽象概念的数学家的想象力.

相同的结构能出现于两个差异极大的数学理论中这个事实引起了关于全部数学的基本统一性的前所未有的不断觉醒，它不顾及基于所研究对象的本性的传统划分. 但是孕育这种数学整体性的观念需要很长的时间，而在 19 世纪，每种新理论的开发大多忽视它同别的理论之间可能的联系. 只是到本章结尾，我们才能概述这种作为当代数学基础及其前进动力的整体性观点(见 § 5，A).

1. 新的演算

A. 复　　数

明显超越古典时代知识的最早的伟大数学发现是解任意的三次方程

$$x^3 - px = q \tag{1}$$

的公式. 附带地我们注意到，任意的三次方程

$$x^3 + a_1x^2 + a_2x + a_3 = 0$$

可表示为上述形式，因为如果令 $x = y - a_1/3$，则我们发现对于 y 所给方程为

$$y^3 - \left(\frac{a_1^2}{3} - a_2\right)y + \frac{2a_1^3}{27} - \frac{a_1a_2}{3} + a_3 = 0.$$

(1)的解可能是波伦亚大学教授希皮奥内·德·费罗于 16 世纪初叶得到的，尽管他并未发表其解法而只是在通信中告诉他的一些弟子和朋友. 他的方法是求形如 $x = u + v$ 的解；我们有

$$x^3 - px = u^3 + v^3 + 3uv(u + v) - p(u + v)$$

$$= u^3 + v^3 + (3uv - p)(u + v).$$

当 $3uv - p = 0$ 时，上式中含有 $u + v$ 的项消失，其结果是如果同时有

$$uv = p/3, \tag{2}$$

$$u^3 + v^3 = q, \tag{3}$$

[115]

那么数 $x = u + v$ 就是(1)的一个解.但由(2)得到

$$u^3 v^3 = \frac{p^3}{27}, \tag{4}$$

而关于二次方程的古典理论表明关系式(3)和(4)蕴涵 u^3 和 v^3 是方程

$$X^2 - qX + \frac{p^3}{27} = 0 \tag{5}$$

的两个根.这就是说，

$$u^3 = \frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}, v^3 = \frac{q}{2} - \sqrt{\frac{q^4}{4} - \frac{p^3}{27}}, \tag{6}$$

从而最后有

$$x = u + v = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}}. \tag{7}$$

(这一公式通常以“卡尔达诺公式”著称，因为卡尔达诺是首次使它为世人所知的学者.)

由于微分学的发明，我们知道如何通过导数来研究函数 $y = x^3 - px - q$ 的变化；由此易于得知方程(1)当 $\frac{q^2}{4} - \frac{p^3}{27} > 0$ 时有 1 个根，当 $\frac{q^2}{4} - \frac{p^3}{27} < 0$ 时有 3 个根.但是，对于第二种情形，公式(7)看起来不再有意义：事实上，根据正负号规则，实数(正或负)的平方总是正的，因此负数不能有平方根.然而，面临这些困难，意大利代数学家及其后继者毫不迟疑地对数 $\sqrt{-a}$ (其中 $a > 0$)进行演算，就好像它们真的存在似的.这就是说，连同规则 $(\sqrt{-a})^2 = -a$，他们对这些东西应用通常的代数规则.例如，

R·邦贝利把卡尔达诺公式应用于方程

$$x^3-15x=4.$$

他知道该方程有一实根 $x=4$,由于

$$x^3-15x-4=(x-4)(x^2+4x+1),$$

该方程还有另外两个实根 $-2\pm\sqrt{3}$.卡尔达诺公式给出的根可写为

[116]

$$x=\sqrt[3]{2+\sqrt{-121}}+\sqrt[3]{2-\sqrt{-121}}. \tag{8}$$

应用二项展开式

$$(a+b)^3=a^3+3a^2b+3ab^2+b^3,$$

得到

$$(2+\sqrt{-1})^3=2^3+3\cdot2^2\sqrt{-1}+3\cdot2(\sqrt{-1})^2+(\sqrt{-1})^3,$$

以 -1 代 $(\sqrt{-1})^2$,以 $-1\cdot\sqrt{-1}$ 代 $(\sqrt{-1})^3$,得到

$$(2+\sqrt{-1})^3=2+11\sqrt{-1}=2+\sqrt{(-1)(11)^2}=2+\sqrt{-121}.$$

他以同样方式算出

$$(2-\sqrt{-1})^3=2-\sqrt{-121},$$

从而由(8)得出

$$x=2+\sqrt{-121}+2-\sqrt{-121}=4.$$

他说,即使有点“矫揉造作”,他觉得这个结果令他满意! 他的先驱卡尔达诺或许是第一位敢于进行此类演算的人,在提及这种演算时说到“道德上的折磨”! 这并不妨碍他们的后继者在二百多年内更大量地不断进行涉及这些他们称之为“不可能的”或“虚的”数的运算.如果 $b>0$,则我们有 $-b=b\times(-1)$,类似于对正数 b,b' 的公式 $\sqrt{bb'}=\sqrt{b}\cdot\sqrt{b'}$,我们写 $\sqrt{-b}=\sqrt{b}\times\sqrt{-1}$.这样所有虚数归结为 $x+y\sqrt{-1}$ 的形式,或者用欧拉引进的对于 $\sqrt{-1}$ 的简略记号 i 写为 $x+iy$,其中 x,y 是实数.坚持使用这种无意义的表示式是基于这样的事实:通过这些运算代数学家发现了获得不必区别若干情形的一般定理的可能性:二次方程恒有一个重根或两个实或虚的单根;实系数三次方程

总至少有一实根①,因而可写为

$$(x-a)(x^2+px+q)=0.$$

如果我们把虚根也算在内,那么三次方程的根的重数之和恒为3.到17世纪初叶时,A.吉拉尔已确信,n 次方程 [117]

$$a_0x^n+a_1x^{n-1}+\cdots+a_n=0(a_0\neq 0)$$

有实或虚根,其重数之和恒为 n.这就是所谓"代数基本定理",但直到18世纪末叶,才由高斯充分证明了这条定理.

到那时候,高斯以及另外两个名气较小的数学家韦塞尔和阿尔冈,终于成功地相互独立地给出涉及数 $a+bi$ 的运算的意义;遵照高斯的说法,这种数称为"复数".为此只须把 $a+bi$ 与关于平面上直角坐标系的坐标为 (a,b) 的点等同起来.稍后由哈密顿提出的一种等价的解释是把 $a+bi$ 看作实数的一个偶

$$\begin{pmatrix} a \\ b \end{pmatrix},$$

并具有对于加法的运算规则

---

① 对 $x\neq 0$,多项式 $x^3+a_1x^2+a_2x+a_3$ 可写为

$$x^3\left(1+\frac{a_1}{x}+\frac{a_2}{x^2}+\frac{a_3}{x^3}\right). \qquad (1)$$

当 $|x|>1$ 时,我们有

$$\left|\frac{a_1}{x}+\frac{a_2}{x^2}+\frac{a_3}{x^3}\right|\leqslant\frac{|a_1|+|a_2|+|a_3|}{|x|},$$

因此

$$1+\frac{a_1}{x}+\frac{a_2}{x^2}+\frac{a_3}{x^3}\geqslant 1-\frac{|a_1|+|a_2|+|a_3|}{|x|}.$$

当 $|x|\geqslant 1$ 和 $|x|>|a_1|+|a_2|+|a_3|=A$ 时,乘积(1)与 x^3 有相同的符号即当 $x<-A-1$ 时为负而当 $x>A+1$ 时为正.用斯蒂文方法(第三章,§6)即可证明方程

$$x^3+a_1x^2+a_2x+a_3=0$$

存在一个实根 a.

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \end{pmatrix} \tag{9}$$

和对于乘法的运算规则

$$\begin{pmatrix} a \\ b \end{pmatrix} \times \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} aa' - bb' \\ ab' + a'b \end{pmatrix}. \tag{10}$$

偶$\begin{pmatrix} a \\ 0 \end{pmatrix}$等同于实数 a①,由此实数作为特殊的复数出现,而规则(10)特别地给出

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}^2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1;$$

如果

$$i = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

这就是“不可能的”运算 $i^2 = -1$.

从 18 世纪以后,数学家们也不加验证地把复数用于分析甚至数论,以得到关于“真实”数的正确结果,这显得相当神奇.——当关于复数的运算得到验证,这些大胆的介入就成为在整个 19 [118] 世纪中发展的一种庞大广博理论的萌芽,这就是由柯西、黎曼、魏尔斯特拉斯和庞加莱开创和发展的复变量解析函数论以及从高斯和狄里克雷到戴德金和希尔伯特开发的代数数论.可惜由于即使着手讨论它们也需要专门的知识,因而我们只能在一般概述(§5,A)中提及这些理论.

B. 向　　量

很早以前,(9)中定义的加法概念就已在别的地方出现.如

① 由于我们有

$$\begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} a + b \\ 0 \end{pmatrix} \text{和} \begin{pmatrix} a \\ 0 \end{pmatrix} \times \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} ab \\ 0 \end{pmatrix},$$

这样的等同是合理的.

果 A 是点$\begin{pmatrix} a \\ b \end{pmatrix}$，$A'$ 是点$\begin{pmatrix} a' \\ b' \end{pmatrix}$，那么

$$\begin{pmatrix} a + a' \\ b + b' \end{pmatrix}$$

是以 OA，OB 为邻边的平行四边形对角线的终点 C①(图 35). 从 17 世纪初叶开始(无疑甚至更早，至少在特殊情形是如此)，这个"平行四边形法则"就已用来确定由两个运动"合成"的运动驱使的点的速度(阿基米德螺线的运动学描述是一个例子，参见第三章，§ 8). 速度按其内涵由一"向量"表示，就是沿着运动方向即运动轨道的切线方向的一个直线段，其长度与速度大小成正比. 罗伯瓦尔特别聪明地运用这个法则来确定某些曲线的切线. 后来牛顿正是通过把力表示为从某个点出发的线段，用同样的办法定义作用于一个点的力的合成. 这种"加法"的性质完全类 [119] 似于实数加法的性质.

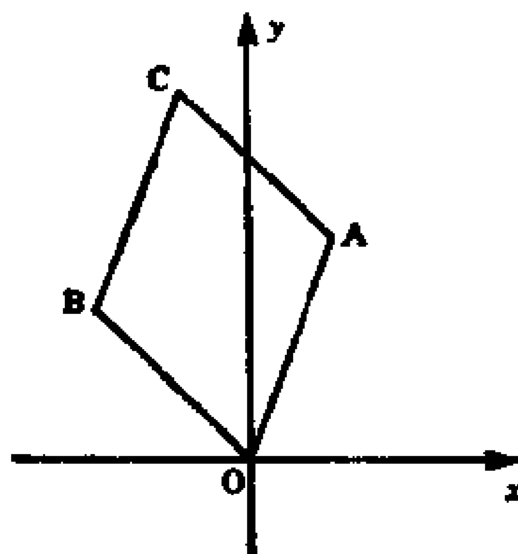


图 35

① 当 O, A, B 在一条直线上时，加法 $C = A + B$ 已在第三章 § 5 中几何地定义.

当然在空间中也能同样构造平行四边形,这就是说,考虑实数构成的三元组

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

并定义其加法为

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \\ c + c' \end{pmatrix}. \tag{11}$$

但是,为什么要在三元组这里停下来? 没有什么理由不能对任意的自然数 n,把 n 个数构成的数组

$$\boldsymbol{x} = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \tag{12}$$

称为 n 维向量,尽管当 $n > 3$ 时,不再能把它与直线段相联系. 数 $a_j(1 \leqslant j \leqslant n)$是向量 $\boldsymbol{x}$ 的分量,而两个向量 $\boldsymbol{x}, \boldsymbol{x}'$ 之和 $\boldsymbol{x} + \boldsymbol{x}'$ 是以 $\boldsymbol{x}$ 与 $\boldsymbol{x}'$ 的分量之和 $a_j + a'_j$ 为分量的向量.

施于向量的另一运算来自平面上的位似概念(见下面的 E),这就是以数 α(为区别于向量,称它为标量)乘向量 $\boldsymbol{x} = \begin{pmatrix} a \\ b \end{pmatrix}$的"标量乘法",它由

$$\alpha \cdot \boldsymbol{x} = \begin{pmatrix} \alpha a \\ \alpha b \end{pmatrix}$$

定义.

这种"乘法"可直接推广:对于等于 n 个数的数组(12)的向量 $\boldsymbol{x}$, $\alpha \cdot \boldsymbol{x}$ 是以 $\alpha a_j(1 \leqslant j \leqslant n)$为分量的向量. 这使人们有可能把 $\boldsymbol{x}$ 表示为如下形式:

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n,$$

其中 e_1 是第一个分量为 1 而其余分量为 0 的向量，e_2 是第二个分量为 1 而其余分量为 0 的向量，以此类推.

在 19 世纪中叶，格拉斯曼还引进了另一种向量运算，它能使坐标的使用大大简化. 这就是平面上两个向量

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

的标量积 $u \cdot v$，它定义为数 [120]

$$u_1 v_1 + u_2 v_2.$$

同时，用这个运算能写出向量 u 的长度和向量 u 与 v 之间夹角 θ 的余弦；前者记为 $\| u \|$，由

$$\| u \| = \sqrt{u \cdot u} \tag{13}$$

给出，后者则由

$$\cos\theta = \frac{u \cdot v}{\| u \| \cdot \| v \|} \tag{14}$$

给出. 由此两个向量的正交性条件可简写为

$$u \cdot v = 0. \tag{15}$$

再者，标量积遵从简单的运算规则：

$$u \cdot v = v \cdot u, \tag{16}$$

$$(u + u') \cdot v = u \cdot v + u' \cdot v, \tag{17}$$

对任一标量 α，
$$(\alpha u) \cdot v = \alpha(u \cdot v). \tag{18}$$

这些性质使得欧几里得的许多命题一旦翻译为向量和标量积语言，就几乎成为自明的. 例如，以 b，c 为直角两夹边之长的直角三角形斜边之长 a 与向量 $u + v$ 的长度相同，这里 u，v 是长度为 b，c 的两个互相垂直的向量(图 36).

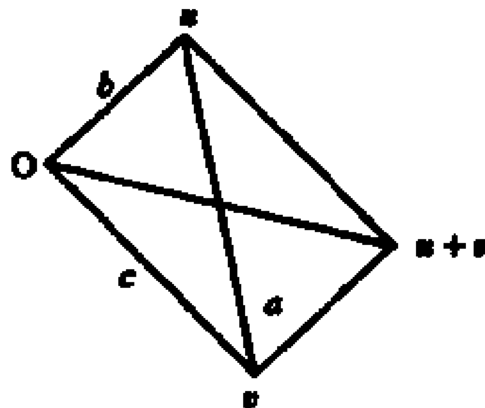


图 36

由(13),(16),(17),我们有

$$a^2 = \| u + v \|^2 = (u + v)\cdot(u + v) = u\cdot u + 2u\cdot v + v\cdot v.$$

由于按假定 $u\cdot v = 0$,我们有

$$\| u + v \|^2 = \| u \|^2 + \| v \|^2,$$

121] 由此得到 $a^2 = b^2 + c^2$,这正是毕达哥拉斯定理.

如果令

$$u\cdot v = u_1v_1 + u_2v_2 + u_3v_3,$$

就能把 $u\cdot v$ 的概念以及前述性质毫无修改地推广到空间中的向量,这里 u_1, u_2, u_3 是向量 u 的三个分量,v_1, v_2, v_3 是向量 v 的三个分量.这样,使用标量积不仅使我们可以离开坐标系,甚至可以不考虑维数.在 20 世纪,当原子物理学中的空间成为"希尔伯特空间"时,这又被重新提起(附录 4,C).

C. 函数的代数运算

我们已看到(第三章,§8),一个实变量的实值函数观念出现于 17 世纪末期.通过作出它的图象这一迂回办法,它可以认为是从人们感官能够感知的平面几何中的"粗略"概念予以"抽象化"而导出的.但从 18 世纪初期起,已有必要考虑几个实变量 $x, y, z, \cdots$ 的(取实值的)函数,例如

$$f(x, y, z) = xy + yz + zx.$$

对于两个自变量,仍有可能谈论函数的"图象",这时它是空间中关于 3 个坐标轴画出的曲面;例如,函数 $f(x, y) = xy$ 的图象是方程为 $z - xy = 0$ 的称为"双曲抛物面"的二次曲面.但只要自变量的个数超过 2,我们就必须放弃任何"形象".在 18 世纪,人们只是说 $f(x, y, z, \cdots)$ 是通过对给定于变量 $x, y, z, \cdots$ 的值进行一定数目的"运算"所得到的值.而且,对于"运算"指的是什么,从未有精确的定义;当然,和、积与商这些代数运算总是提到的,但常常加之以指数 x^y,对数,三角函数以及方程的根(甚至并未解出),等等.函数观念慢慢扩展,最后由狄里克雷给出一般

定义:这个概念必须承担的全部内容在于,相应于诸自变量所取的每组值,有某种方法得到一个一意确定的实数.狄利克雷给出了一个迄今仍然著名的例子,即由下述条件定义的函数 $f(x)$:如果 x 是有理数,则令 $f(x)=0$;如果 x 是无理数,则令 $f(x)=1$.显然不可能画出这个函数的"图象".

从约翰·伯努利起,就用记号 φx 或 $f(x)$ 表示一个没有特别规定的函数在自变量值为 x 时所取的值.从 19 世纪初期起,就逐渐习惯于把函数(至少对于单元函数)看作单个的对象,而不是一连串的值,并把它简记为 f,例如 sin 或 log.对于这些对象也进行运算:显然 $f+g$ 与 fg 由 [122]

$$(f+g)(x)=f(x)+g(x),(fg)(x)=f(x)g(x)$$

定义.

但是对函数还有别的运算:两个函数 f 与 g 的复合,它使 $f(g(x))$ 对应于 x,如今我们把它写为 $f\circ g$ 以与乘积 fg 相区别①;反函数 f^{-1} 当它有意义时②,意味着 $f^{-1}(f(x))=x$, $f(f^{-1}(y))=y$.例如,定义于区间[0,1]上的函数

$$\sin(\pi x/2)$$

取值于同一区间[0,1],它在该区间上具有"反函数"

$$\frac{2}{\pi}\arcsin x.$$

D. 排列和置换

从 16 世纪以来涉及概率计算的问题以及 18 世纪行列式理

① 自然,如果 f 定义在区间 I 上而 g 定义在区间 J 上,则必须对 J 中的每个数 x, $g(x)$ 属于 I.

② 当然必须是这样的情形:对两个不同的数 x,x',不能有 $f(x)=f(x')$.此时对 f^{-1} 的图象有一简单的解释:此图象可通过 f 的图象关于直线 $y=x$(称为"第一分角线")作对称得到.

论中出现的"组合论"问题,导致考虑有限多个对象有序排出的各种不同方法.这些排出称为这些对象的排列.例如,对3个对象 a,b,c,共有6种排列:

$$abc, acb, bac, bca, cab, cba.$$

一般地,对于 n 个不同的对象 $x_1, x_2, \cdots, x_n$,易于通过对 n 的归纳推理证明它们的排列总数是乘积

$$1\cdot 2\cdot 3\cdot \cdots \cdot n,$$

它写为 $n!$ 并称作"n 阶乘"(附录2,A).

当 $x_1, x_2, \cdots, x_n$ 是互不相同的(实或复)数时,易于构造一个 n 次代数方程,它以这 n 个数为其根,这就是

$$P(t) \equiv (t-x_1)(t-x_2)\cdots(t-x_n)=0.$$

如果我们写

[123]
$$P(t)=t^n-a_1t^{n-1}+a_2t^{n-2}-\cdots+(-1)^n a_n, \tag{19}$$

那么显然当因子 $(t-x_1), (t-x_2), \cdots, (t-x_n)$ 以不同顺序排列时,系数 $a_1, a_2, \cdots, a_n$ 不会改变.这些系数是 $x_1, x_2, \cdots, x_n$ 的多项式,当序列

$$x_1, x_2, \cdots, x_n$$

代之以它的任一排列时,这些多项式不会有改变.这易于通过直接计算来验证,因为直接计算给出

$$a_1 = x_1 + x_2 + \cdots + x_n,$$

$$a_2 = x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n,$$

$$\cdots\cdots$$

$$a_n = x_1x_2\cdots x_n.$$

一般地,a_k 是从数 $x_1, x_2, \cdots, x_n$ 中以各种可能方式取 k 个并予以相乘得到的所有乘积之和.这是因为,为求出乘积 $P(t)$,只须以各种可能的方式在每个因子 $t-x_1, t-x_2, \cdots, t-x_n$ 中取一项,并把这些项的所有乘积加在一起.

从17世纪起就已知道,每个对称多项式 $S(x_1, x_2, \cdots, x_n)$

(即在 $x_1, x_2, \cdots, x_n$ 的任何排列下都不改变的多项式)都能写成 $R(a_1, a_2, \cdots, a_n)$ 的形式;$R(a_1, a_2, \cdots, a_n)$ 是 $a_1, a_2, \cdots, a_n$ 的有理系数多项式,如果 $S(x_1, x_2, \cdots, x_n)$ 的系数是有理数.例如,对于 $n=3$,

$$x_1^2+x_2^2+x_3^2=(x_1+x_2+x_3)^2-2(x_1x_2+x_2x_3+x_3x_1)$$
$$=a_1^2-2a_2,$$
$$x_1^2x_2+x_1^2x_3+x_2^2x_1+x_2^2x_3+x_3^2x_1+x_3^2x_2$$
$$=(x_1+x_2+x_3)(x_1x_2+x_2x_3+x_3x_1)-3x_1x_2x_3$$
$$=a_1a_2-3a_3.$$

华林于 1770 年在 $S(x_1, x_2, \cdots, x_n)$ 为已知的情形下给出 $R(a_1, a_2, \cdots, a_n)$ 的明显公式.

我们已提到(见上面的 A)希皮奥内·德·费罗关于解三次方程的令人激动的发现.大约在 1540 年,卡尔达诺的学生 L·费拉里,通过复杂得多的公式类似地得到了四次方程的解法;如同公式(7),他在其公式中用到根号的叠置.然而 200 多年中最优秀的数学家们为得到解五次方程的类似公式所作的全部努力都告失败.

1770 年,拉格朗日决定通过把这个问题同方程的根的排列相联系,来寻求总是重蹈覆辙的原因.解四次方程的所有途径都是通过预先解一个辅助三次方程达到的;这使拉格朗日想到用 [124] 下面的观察作为起点来解释这个事实:如果 x_1, x_2, x_3, x_4 是方程

$$t^4-a_1t^3+a_2t^2-a_3t+a_4=0$$

的根,则多项式

$$x_1x_2+x_3x_4$$

不是对称的,但当 x_1, x_2, x_3, x_4 以 $4!=24$ 种不同方式排列时,我们得到的不是 24 个不同多项式,而只是 3 个不同的多项式,即

$$s_1 = x_1x_2 + x_3x_4, s_2 = x_1x_3 + x_2x_4, s_3 = x_1x_4 + x_2x_3.$$

初等对称函数

$$b_1 = s_1 + s_2 + s_3,\quad b_2 = s_1s_2 + s_2s_3 + s_3s_1, b_3 = s_1s_2s_3$$

本身是 x_1, x_2, x_3, x_4 的对称多项式，从而是 a_1, a_2, a_3, a_4 的多项式，而 s_1, s_2, s_3 是辅助三次方程

$$t^3 - b_1t^2 + b_2t - b_3 = 0$$

的根，它可由希皮奥内·德·费罗公式表示. 由此出发，最终易于通过 3 个二次方程的解导出 x_1, x_2, x_3, x_4（附录 1）.

一当这样导出四次方程的解，拉格朗日由此立即作出下面的结论：如果人们能够类似地求出有关五次方程 5 个根的一个多项式 $F(x_1, x_2, x_3, x_4, x_5)$，它当这 5 个根以 $5! = 120$ 种方式排列时只取三或四个值，那么就能"用根式"解五次方程. 但是他怀疑这样的多项式的存在性.

19 世纪初，拉格朗日的学生、意大利人 P·鲁菲尼有效地证明这样的多项式不能存在. 通过我们不能在这里概述的很长的论证，他研究 x_1, x_2, x_3, x_4, x_5 的使给定的多项式 $P(x_1, x_2, x_3, x_4, x_5)$ 保持不变的所有排列形成的系. 他证明该系中排列的数目 p 整除 120，而把 x_j 的 120 种排列施于 P 所能形成的不同多项式的数目是 $120/p$. 最后他证明 $120/p$ 决不可能等于 3 或 4. 做到这一点还未能令人信服地证明任意五次多项式不可能通过根式叠置求解，缺乏关于很久后被称为数域（见 § 3, C）的想法妨碍鲁菲尼达到上述结论；而阿贝尔于 1824 年由于拥有这种想法而能完成这个结论的证明.

这里对我们而言重要的事情在于，晚于鲁菲尼若干年，年青的柯西为推广这些结果，从崭新的角度提出排列概念，而后来的[125]进展表明这具有决定性意义. 例如，如果我们取 4 个对象并排成 a, b, c, d 的次序，又如果 c, b, d, a 是四者的一个排列，那么柯西对此排列建立了下述规则

$$\begin{matrix} a & b & c & d \\ \downarrow & \downarrow & \downarrow & \downarrow \\ c & b & d & a \end{matrix},$$

它使处于第一行的每个对象对应到处于第二行的对象.他称这个规则为置换,并把它写为

$$\begin{pmatrix} a & b & c & d \\ c & b & d & a \end{pmatrix}.$$

而且他理解,4 个对象排列的次序对置换的定义毫不重要,例如

$$\begin{pmatrix} a & b & c & d \\ c & b & d & a \end{pmatrix}$$

与

$$\begin{pmatrix} c & b & d & a \\ d & b & a & c \end{pmatrix}$$

是相同的置换(参看图 37).

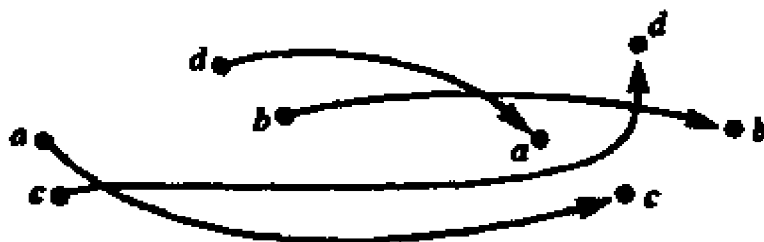


图 37

柯西并没有限于 4 或 5 个对象,而是着手讨论任意数目个对象的情形.虽然他对任意的置换提出了一般记号

$$\begin{pmatrix} a & b & c & \cdots & l \\ \alpha & \beta & \gamma & \cdots & \lambda \end{pmatrix},$$

但很清楚这类"形象"只适用于对象数目很小的情形,因而柯西在作出他的论证时用了简略记号

$$\binom{A}{B},$$

这里 A,B 是所涉及的对象的任一排列.

柯西最重要的创新是两个置换的复合:对于先做置换$\binom{A}{B}$,再做置换$\binom{C}{D}$所得到的置换,他写为

$$\binom{A}{B}\binom{C}{D}.$$

[126] 由于存在某个排列 E,使得$\binom{C}{D}$也能写成$\binom{B}{E}$,所以我们有

$$\binom{A}{B}\binom{B}{E}=\binom{A}{E};$$

例如,

$$\begin{pmatrix} a & b & c & d \\ b & a & c & d \end{pmatrix}\begin{pmatrix} b & d & c & a \\ d & a & b & c \end{pmatrix}$$
$$=\begin{pmatrix} a & b & c & d \\ b & a & c & d \end{pmatrix}\begin{pmatrix} b & a & c & d \\ d & c & b & a \end{pmatrix}$$
$$=\begin{pmatrix} a & b & c & d \\ d & c & b & a \end{pmatrix}.$$

柯西还考虑几个置换的复合,如

$$\binom{A}{B}\binom{C}{D}\binom{E}{F}\binom{G}{H}$$

等,他特别地用记号

$$\binom{A}{B}^k$$

表示 k 个都等于$\binom{A}{B}$的置换的复合.对"恒等置换"即保持每个对象不变的置换,他写为

$$\begin{pmatrix} A \\ A \end{pmatrix}.$$

虽然没有明说，但柯西不会注意不到

$$\begin{pmatrix} A \\ B \end{pmatrix}\begin{pmatrix} B \\ A \end{pmatrix} = \begin{pmatrix} A \\ A \end{pmatrix}.$$

为表明恒存在自然数 N，使得

$$\begin{pmatrix} A \\ B \end{pmatrix}^N = \begin{pmatrix} A \\ A \end{pmatrix},$$

我们论证如下：由于只有有限个排列，因此当不断作置换

$$\begin{pmatrix} A \\ B \end{pmatrix}^m$$

时，必定有两个不同的自然数 h,k，使得

$$\begin{pmatrix} A \\ B \end{pmatrix}^h = \begin{pmatrix} A \\ B \end{pmatrix}^k,$$

我们可设 $h>k$. 由上得

$$\begin{pmatrix} A \\ B \end{pmatrix}^h\begin{pmatrix} B \\ A \end{pmatrix}^k = \begin{pmatrix} A \\ B \end{pmatrix}^k\begin{pmatrix} B \\ A \end{pmatrix}^k.$$

但因

$$\begin{pmatrix} A \\ B \end{pmatrix}\begin{pmatrix} B \\ A \end{pmatrix}$$

是恒等置换，故可写

$$\begin{pmatrix} A \\ B \end{pmatrix}^p\begin{pmatrix} B \\ A \end{pmatrix}^q = \begin{pmatrix} A \\ B \end{pmatrix}^{p-1}\begin{pmatrix} A \\ B \end{pmatrix}\begin{pmatrix} B \\ A \end{pmatrix}\begin{pmatrix} B \\ A \end{pmatrix}^{q-1} = \begin{pmatrix} A \\ B \end{pmatrix}^{p-1}\begin{pmatrix} B \\ A \end{pmatrix}^{q-1},$$

因此我们断定

$$\begin{pmatrix} A \\ B \end{pmatrix}^{h-k} = \begin{pmatrix} A \\ A \end{pmatrix}.$$

对于我们，上述复合与函数复合（见 C）的类似性立即跳进心头．但对于 19 世纪初叶的数学家，一个有限集合与一条直线之间差别太大，难于相信两者可以统一起来；这种统一只能随着

戴德金和弗雷格才能到来.

E. 位移和仿射变换

显然平面上或空间中刚性图形的运动这种想法是每个人都[127]懂的事情,但古希腊人是否有位移概念,仍不能肯定.位移是指一个图形从其最初位置转移到最终位置而不考虑其中间位置.欧几里得所说的两个三角形的"相等"(在希尔伯特所称的"全等"意义下)①,最初无疑意味着作一个运动使一个三角形置于另一三角形之上的可能性②.但是把位移本身看作独立于它所移动的"图形"的数学对象,却是18世纪之前几乎从未出现过的思想方法③.

那时欧拉把位移观念解释为"整体地"应用于整个空间的一个"变换",即使得空间的每个点(不仅是一个"图形"中的点)与另一个点相联系的一个法则.两个这样的变换显然可以通过先作第一个再作另一个来"复合".欧拉证明,空间中每个使某个点不动的位移是绕通过该点的某条直线的旋转,而每个位移可由一个旋转与一个平移复合得到.在附有两互相垂直的坐标轴的平面上,每个位移可通过使每个点(x, y)对应到坐标为

---

① 欧几里得用"相等"一词表示几种意义:有时指两个能叠置的图形,有时指两个具有相同面积的图形.

② 欧几里得在用这种方法"证明"两种情形下三角形相等(《几何原本》第Ⅰ卷命题4,8)时简单地说一个三角形"贴于"另一个之上(εφαρμόζεσθαι).这是他如何引出他没有定义过且没有以此作为任何公设的主题的概念的又一个例子.在16世纪,欧几里得的评注者之一J·皮莱蒂埃对这种"生硬插进"的办法感到懊恼,认为这对"几何学的神圣性"带来怀疑,而且已经想到欧几里得应当把这些命题取作公设(而这正是希尔伯特对其中之一所作的处理).在后面各卷中,欧几里得甚至在这种办法本来很自然时,也避免求助于这种论证方法.

③ 然而在笛卡儿那里我们确实发现"无穷小旋转"的想法.

$$X = a + x\cos\theta - y\sin\theta, Y = b + x\sin\theta + y\cos\theta \quad (20)$$

的点来表示,其中 a,b 和 θ 是刻画该位移的实数.更一般地,对于类似地由

$$X = a + \alpha x + \beta y, Y = b + \gamma x + \delta y \quad (21)$$

定义的变换,则用仿射变换这个名称,这里 $a,b,\alpha,\beta,\gamma,\delta$ 是任意的实数.欧拉已想到特殊情形

$$X = \alpha x, Y = \beta y,$$

而 19 世纪的几何学家广泛使用的位似变换是 $\alpha = \beta$ 的情形,其具体"形象"可借助比例绘图器画出,它是绘图员用来保持固定 [128] 比例放大或缩小平面图形的器具(图 38).此仪器的杆可围绕 O 旋转,并在 A,B,A',B' 处以铰链连接,形成两个边长为 a,a' 的菱形;点 M,M' 与 O 处于一直线上,且 $\overline{OM'}/\overline{OM} = -a'/a$,这样当 M 描绘曲线 C 时,M' 描画出比例为 a'/a 的位似曲线.

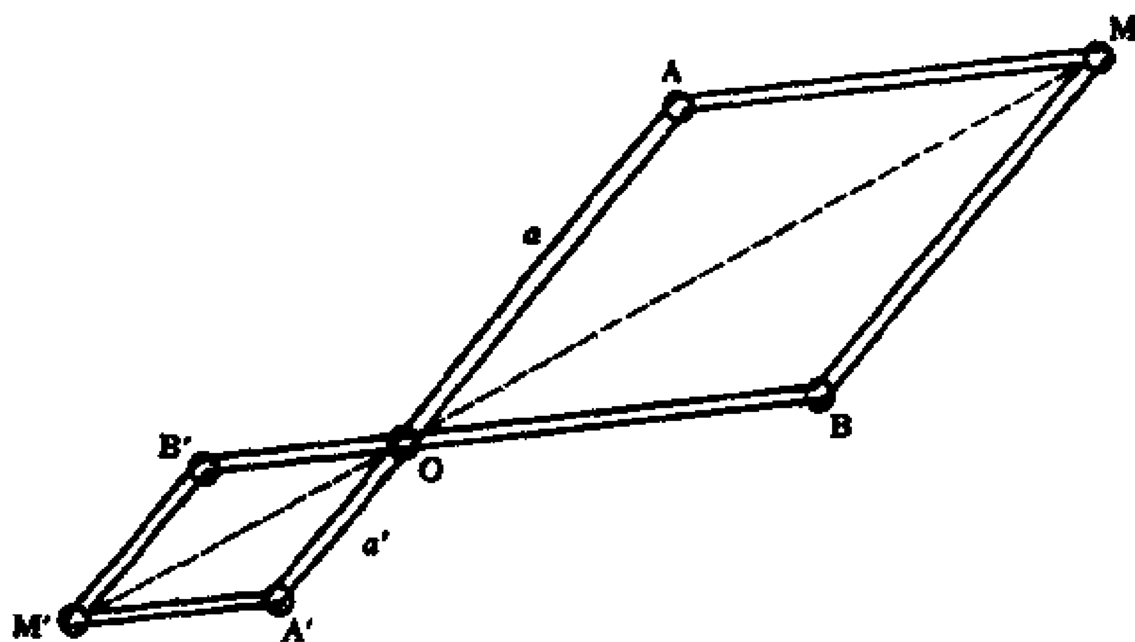


图 38

F. 整数同余式的演算

许多算术问题(例如参见第四章附录 3)必须处置能被给定

整数 $m>1$ 整除的(正、负或零)整数即形如 km 的整数,其中 k 是整数.更一般地还考虑整数 $a+km$,其中 a,m 固定而 k 是任一整数(可以说这是"沿两个方向的"算术序列).从高斯时代以来,用

$$x \equiv a(\mathrm{mod}\, m) \tag{22}$$

来代替"x 形如 $a+km$"已成习惯,式(22)读作"x 模 m 同余于 a",并说此关系式是整数模 m 的同余式.

这个记号显然使人想起相等关系 $x=a$.这种记号相似性是完全合理的:事实上,从两个关系式

$$x \equiv a(\mathrm{mod}\, m),\ y \equiv b(\mathrm{mod}\, m) \tag{23}$$

可推出

[129]

$$x+y \equiv a+b(\mathrm{mod}\, m) \tag{24}$$

和

$$xy \equiv ab(\mathrm{mod}\, m), \tag{25}$$

因为对 $x=a+hm, y=b+km$,只须算一下即有

$$x+y=a+b+(k+h)m,\ xy=ab+(hb+ka+hkm)m.$$

于是对任一自然数 $m>1$,我们定义了两种新的"合成律".对于每个整数 x,有且只有一个自然数 r 满足 $0\leqslant r\leqslant m-1$ 且 $x\equiv r(\mathrm{mod}\, m)$.我们说 r 是数 $r+km$ 构成的"类"$\bar{r}$ 的"代表"(这是欧拉已经使用的术语).然而,当

$$0\leqslant r\leqslant m-1, 0\leqslant s\leqslant m-1$$

时,自然可有 $r+s>m$ 或 $rs>m$;$\overline{r+s}$(或 $\overline{rs}$)的"代表"是 $r+s$(或 rs)除以 m 的余数.特别地,即使 $r\not\equiv 0(\mathrm{mod}\, m)$,$s\not\equiv 0(\mathrm{mod}\, m)$,也可能有 $rs\equiv 0(\mathrm{mod}\, m)$.例如对 $m=6$ 取 $r=2, s=3$.此时我们说类 $\bar{r}$ 和 $\bar{s}$ 是关于同余式演算的零因子.

G. 二次型类的演算

代数恒等式

$$(a^2+b^2)(c^2+d^2)=(ac-bd)^2+(ad+bc)^2 \tag{26}$$

无疑可追溯到古典时代. 拉格朗日使用更一般的恒等式

$$(x^2+Dy^2)(u^2+Dv^2)=(xu-yvD)^2+D(xv+yu)^2, \quad (27)$$

实际上此恒等式起源于印度数学家波罗摩笈多(7 世纪). 拉格朗日由此推出,如果方程

$$x^2+Dy^2=n_1,$$
$$u^2+Dv^2=n_2$$

(其中 D 是整数)分别具有整数解(x,y),(u,v)(第四章, §2, A),则方程

$$X^2+DY^2=n_1n_2$$

具有整数解.

高斯发现一个比(27)难记得多的规律. 一般地说,我们不能有形如

$$(ax^2+bxy+cy^2)(au^2+buv+cv^2)=aX^2+bXY+cY^2$$

的恒等式,其中 X,Y 关于 x,y 和关于 u,v 均为一次多项式:

$$\begin{cases}X=p'xu+q'yv+r'xv+s'yu,\\ Y=p''xu+q''yv+r''xv+s''yu,\end{cases} \quad (28)$$

而系数 $p',q',\cdots,r'',s''$ 都是整数. 然而,通过非常繁复的计算, [130] 高斯证明,如果 $ax^2+bxy+cy^2$,$a'x^2+b'xy+c'y^2$ 是两个整系数二元二次型,并具有相同的判别式 Δ,则恒存在系数为 $\alpha,\beta,\gamma,\delta$ 且满足 $\alpha\delta-\beta\gamma=1$ 的变量替换

$$\begin{cases}z=\alpha x+\beta y,\\ t=\gamma x+\delta y,\end{cases}$$

它具有下述性质:如果 $az^2+bzt+ct^2=Ax^2+Bxy+Cy^2$,则有恒等式

$$(Ax^2+Bxy+Cy^2)(a'u^2+b'uv+c'v^2)$$
$$=A'X^2+B'XY+C'Y^2,$$

其中 X,Y 由整系数公式(28)给出. 此外,二次型 $A'X^2+B'XY+C'Y^2$ 也具有判别式 Δ,且按此方法得到的所有二次型是等价的,它们的类(第四章, §2, A) C'' 只依赖于二次型 ax^2+bxy+

cy^2 和 $a'x^2+b'xy+c'y^2$ 的类 C 和 C'.然后高斯称类 C'' 由类 C 和 C' 合成并写为 $C''=C+C'$.

2. 第一种结构

A. 合成律的主要性质

§1 中所给的例子都涉及通称的合成律,它可概括地描述为一种手段,它能实现属于同一类的两个对象 a,b 与该类的第三个一意确定的对象 c 之间的一个对应;这个对应可写为 $c=a\bowtie b$,其中 $\bowtie$ 是刻画所考虑的规律的记号.最常用的记号是 $+$,$\cdot$,$\times$,$\circ$,$*$,$\cup$,$\cap$,$\wedge$ 和 $\otimes$.

欧几里得早已想到(《几何原本》第Ⅶ卷,命题 16),对于两个自然数 a 与 b 的乘积,$a\times b=b\times a$ 这个事实需要证明.这就是到 19 世纪初叶时所称的乘法的交换性,面对当时表述的新的合成律,这个问题立即被提了出来.显然对于柯西的"置换"(§1,D),一般没有交换性.换言之,如果我们(如同伽罗瓦那[131]样)用一个单个字母表示置换,则完全可能有 $TS\neq ST$.例如,如果取

$$S=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},T=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},$$

那么情形就是如此,因为我们有

$$ST=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},TS=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}.$$

对两个实变函数的合成即复合,情形相同.例如,如果 $f(x)=x^2$,$g(x)=x+1$,那么

$$f(g(x))=(x+1)^2,g(f(x))=x^2+1.$$

平面上位移的合成(即复合)也不是交换的.平移 $T:(x,y)\longmapsto(x+1,y)$与旋转 $R:(x,y)\longmapsto(-y,x)$不相交换.

另一方面,在数的同余式演算(§1,F)中,我们有

$$a+b\equiv b+a \pmod m,$$
$$ab\equiv ba \pmod m.$$

同样，高斯证明，对具有相同判别式的两个二次型类 C，C'，有 $C+C'=C'+C$.

合成律的另一性质是所谓中性元的存在性：如果把合成律记为 $a \bowtie b$，那么中性元就是一个元素 e，它对每个元素 a 满足
$$e \bowtie a = a \bowtie e = a.$$
如果这样的元素存在，那么它是唯一的. 最显见的例子是对于实数加法的 0 和对于实数乘法的 1. 对于实变函数，"恒等"函数 $e(x)=x$ 是合成律 $f\circ g$ 的中性元. 同样，对于置换（或平面上的位移），使所有对象（或所有点）不动的"恒等"置换（或位移）是中性元. 对于具有判别式 Δ 的二次型类的合成，中性元的存在性很不显见；高斯证明，如果 $\Delta\equiv 0 \pmod 4$，则中性元 E 是二次型
$$x^2-\frac{1}{4}\Delta y^2$$
的类，如果 $\Delta\equiv 1 \pmod 4$，则 E 是二次型
$$x^2+xy-\frac{1}{4}(\Delta-1)y^2$$
的类.

此外，中性元的存在性并非总是有保证的. 例如，两个偶整数之积是偶整数，但对于偶整数的乘法没有中性元. [132]

当中性元存在时，我们可以问是否存在元素 a 的"逆元" a'，即使得
$$a \bowtie a' = a' \bowtie a = e$$
的元素 a'. 很早就已知道分数 m/n 对于乘法具有"逆" n/m. 引进负数使得每个正整数 n 对于加法具有"逆" $-n$ 成为可能，但在这里用的是"负"这个词. 我们也已注意到按柯西所用记号，置换 $\begin{pmatrix}A\\B\end{pmatrix}$ 具有对于复合的"逆"置换 $\begin{pmatrix}B\\A\end{pmatrix}$；同样平面上每个位移有其"逆"，即使得每个点返回到最初位置的位移. 另一方面，对于

定义在 $\mathbf{R}$ 上的实函数,例如函数 $f(x)=x^3$ 关于函数复合有“逆”即其反函数 $g(x)=x^{1/3}$,但对 $f(x)=x^2$ 此点不成立,因为不可能有实函数 g,使当 x 为负时有 $(g(x))^2=x$. 对于数的同余式(§1,F),关于它的乘法不一定都有逆. 例如,不存在自然数 n 使得 $2n\equiv1(\mathrm{mod}4)$,因此 2“模 4”没有逆(附录 3,A). 另一方面,高斯证明每个二次型类 C 有一“逆”C' 使得 $C+C'=E$. (我们仍喜欢说 C' 是 C 的“相反类”;当合成律记为 + 时,通常用“相反”这个词代替“逆”).

还有一个性质在 19 世纪初叶之前一直没有被明确指出来,这就是合成律的结合性. 用一般记号 $\bowtie$,它可写为

$$(a\bowtie b)\bowtie c=a\bowtie(b\bowtie c).$$

无疑这是由于该性质对所有前面提及的已知合成律都成立. 对于实数的加法和乘法,它被不自觉地运用,而对函数(或置换、位移)的复合,它恰恰也很“显然”,因为 $(f\circ g)\circ h$ 和 $f\circ(g\circ h)$ 两者都是在 x 处取值为 $f(g(h(x)))$ 的函数. 需要证明的唯一情形是二次型类的合成,这是由高斯明确提供的. 直到 1845 年,才定义非结合的合成律.

最后,当相同的一些对象有两种合成律时,就出现分配性;例如,数的乘法关于数的加法的分配性

[133]

$$a(b+c)=ab+ac.$$

这个性质也是在 19 世纪初才被明显注意到的. 然而它并不总是满足的. 例如实函数的复合关于其加法确实有

$$(f+g)\circ h=(f\circ h)+(g\circ h),$$

但另一方面一般地说

$$f\circ(g+h)\neq(f\circ g)+(f\circ h).$$

B. 变换群

我们看到(§1,D),柯西于 1815 年引进了对任一有限对象

集定义的两个“置换”的合成 ST.如他的论文的标题所指出的[①],他关心的中心问题是拉格朗日和鲁菲尼已经开始讨论的问题:确定 n 个变量 $x_1,x_2,\cdots,x_n$ 的使得给定多项式 $P(x_1,x_2,\cdots,x_n)$ 不变的置换.很显然(不管我们有何种多项式 P),如果 S,T 是具有上述性质的两个置换,则 ST 和 S^{-1} 也同样具有这个性质.此外,如果 U 是把 P 变换为一个不同的多项式 Q 的置换[②],则由定义,如果 S 使 P 不变,那么 SU 也把 P 变换为 Q.反之,如果置换 V 也把 P 变换为 Q,则 $S=VU^{-1}$ 使 P 不变,因而 V 能写为 $V=V(U^{-1}U)=(VU^{-1})U=SU$.柯西考虑由所有 $n!$ 个可能的置换得到的所有不同的多项式 $P=P_1,P_2,\cdots,P_d$;令 $U_1=I,U_2,\cdots,U_d$ 是分别把 P 变换到 $P_1=P,P_2,\cdots,P_d$ 的置换,另一方面,令 $S_1=I,S_2,\cdots,S_m$ 是所有使 P 不变的置换.然后以下列方式把所有 $n!$ 个置换划分成一些类:

$U_1=I=S_1,S_2,S_3,\cdots,S_m$ 使 P 不变,

$U_2,S_2U_2,\cdots,S_mU_2$ 把 P 变换为 P_2,

……

$U_d,S_2U_d,\cdots,S_mU_d$ 把 P 变换为 P_d.

由于每一行都有相同的项数 m,因此有

$$n! = md; \tag{29}$$ [134]

除术语外,这事实上是拉格朗日已得出的结果.柯西正是从这个结论继续前进,通过精巧的置换组合,证明如果 n 是素数,则必

① Mémoire sur le nombre des valeurs qu'une fonction peut acquérir lorsqu'on y permute de tous les manières possible les quantités qu'elle renferme(《关于一个函数其所含变量以各种可能方式置换时能取之值的个数》).

② 例如,置换

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_4 & x_1 & x_3 \end{pmatrix}$$

把多项式 $x_1x_2+x_3x_4$ 变为 $x_2x_4+x_1x_3$.

有 $d=2$ 或 $d \geqslant n$，从而推广了鲁菲尼的结果.

抽象化的另一步是伽罗瓦于 1830 年跨出的.他当时不是继续谈论不变多项式，而是对 $n!$ 个置换构成的集合的一个子集 Γ 使用 n 个对象的置换群①这一术语，这里 Γ 满足：如果 S，T 在 Γ 中，那么它们的"积"ST 也在 Γ 中.由于这特别地意味着对每个指数 k，S^k 在 Γ 中，我们能写 $SS^{k-1}=S^{k-1}S=I$，这里设 $S^k=I$.由此推出 S^{k-1} 是 S 的逆 S^{-1}，并且 S^{-1} 也在 Γ 中.$n!$ 个置换的全体 $\mathfrak{S}_n$ 显然是一个群，称为 n 个对象的对称群.如果 Γ 是一个置换群，Γ' 是另一个包含于 Γ 中的群，则称 Γ' 是 Γ 的一个子群.按定义，置换群的阶是它的元素的个数.如果 Γ' 是 Γ 的子群，g，g' 分别是 Γ，Γ' 的阶，则 g' 是 g 的因数.

如果 S 是任一置换，则存在使得 $S^h=I$ 的最小的数 h；于是置换 $I, S, S^2, \cdots, S^{h-1}$ 互不相同，因为如果有 $S^p=S^q$，$1 \leqslant p<q<h$，我们就会推出 $(S^{-1})^p S^p=(S^{-1})^p S^q$，就是说 $I=S^{q-p}$，$q-p<h$，这与 h 的定义矛盾.这 h 个置换构成一个子群，因为 $S^p S^q=S^{p+q}$ 而如果 $h \leqslant p+q<2h$，则我们有 $S^{p+q}=S^h S^{p+q-h}=S^{p+q-h}$.我们称这样的子群是循环的且它由 S 生成，我们也称 h 是 S 的阶.

循环群的一个例子是由通称为循环置换的

$$S=\begin{pmatrix}1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1\end{pmatrix}$$

的幂构成的群.如果把置于正 n 边形顶点上的数 $1,2,\cdots,n$ 放在一个圆上，则 S 可用一个旋转来表示，它使每个顶点转至下一个顶点，而最后一个顶点则转至第一个顶点(图 39).由此立即得知，对于 $1 \leqslant p<n$，我们有

① 柯西对"置换"用的是"substitution"这个词，对此伽罗瓦有时用"permutation".后来对"置换群"流行用"group of permutations"这个术语.

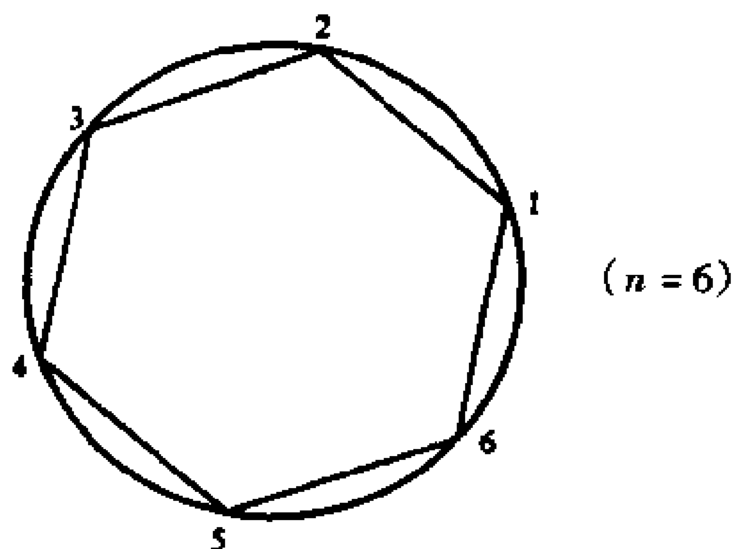


图 39

$$S^p=\begin{pmatrix}1 & 2 & 3 & \cdots & n-1 & n\\ p+1 & p+2 & p+3 & & p-1 & p\end{pmatrix},$$

[135]

因此置换 $S^p(1\leqslant p<n-1)$互不相同且 $S^n=I$. 这个置换可简写为 $S=(1\quad 2\quad 3\quad \cdots\quad n)$.

子群的例

使用前面的记号,对称群 $\mathfrak{S}_3$ 的 6 个元素可写为

$$I,(1\quad 2),(1\quad 3),(2\quad 3),(1\quad 2\quad 3),(1\quad 2\quad 3)^2,$$

其中

$$(1\quad 2\quad 3)^2=\begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix}.$$

它有由

$$I,(1\quad 2\quad 3),(1\quad 2\quad 3)^2$$

构成的一个 3 阶循环子群 $\mathfrak{A}_3$. 它还有 3 个 2 阶子群

$$\{I,(1\quad 2)\},\{I,(1\quad 3)\},\{I,(2\quad 3)\}.$$

最后,它有一个仅由中性元 I 构成的子群.

正是这些概念及其极富独创性的运用,使伽罗瓦得以通过证明"用根式"解代数方程问题等价于关于现今称为所给方程的

"伽罗瓦群"的一个问题(附录 2,B),从而完全解决了前一问题.也正是伽罗瓦理论及群论的发展使得决定所有"能用直尺和圆规"作图的几何问题成为可能.需要特别强调的是,为达到这些问题的解,必须坚决地抛开经典代数并引进新的"抽象"对象即置换群,而后者与代数方程有着完全不同的本性.后来黎曼和狄里克雷通过引进解析函数研究素数(第四章,§2,B),也是继承同样的办法.这必须看作数学同传统划分决裂的深刻统一性的[136] 最早的清晰显示.

伽罗瓦去世若干年后,晶体学家开始对看来远离方程论的一个几何问题感到兴趣:确定空间中使一个给定的多面体保持不变的旋转.例如,在立方体情形,这些旋转必须使立方体的中心不动,因此旋转的轴都应通过立方体中心.此外,这些旋转使立方体的 8 个顶点作一排列,而且两个这样的旋转之积还是一个使此立方体不变的旋转.这样这些旋转构成 8 个顶点的排列的群 $\mathfrak{S}_8$ 的一个子群 Γ.事实上,证实 Γ 的阶为 24(而 $\mathfrak{S}_8$ 的阶为 $8! = 40320$)并明显地定出 Γ 中两个旋转的积,是一个很好的练习(附录 2,E).不过这些群同柯西和伽罗瓦的变换群的相似性起初并未引人注意,只是很晚以后晶体学家们才使用群的语言.

平面上和空间中的几何学在 19 世纪前半期取得了巨大的进步.在这个领域里,拉格朗日和高斯的"变量替换"(第四章,§2,A)终于被解释为"变换".它们是例如下述替换的特殊情形:这种替换实现平面上坐标为(x,y)的点与坐标为(X,Y)的点之间由

$$S:\begin{cases} X = a + \alpha x + \beta y, \\ Y = b + \gamma x + \delta y \end{cases} \tag{30}$$

给定的对应,其中 $a,b,\alpha,\beta,\gamma,\delta$ 是任意的实数.这就是所谓平面"仿射变换"(§1,E).仿效高斯和艾森斯坦,我们用一个单个字母表示这样的变换.两个这样的变换的复合(或"积")ST 还是

这样的变换,而当 $\alpha\delta-\beta\gamma\neq0$ 时,(30)能对 x,y 解出,从而给出一个仿射变换 S^{-1},它是 S 的"逆".然而,在 1860 年之前,这些性质几乎从未被表述为:满足 $\alpha\delta-\beta\gamma\neq0$ 的变换(30)构成一个群,即仿射群 $A(2)$.由公式(20)定义的位移组成上述群的一个子群.另一个重要子群是由满足 $a=b=0,\alpha\delta-\beta\gamma\neq0$ 的变换(30)构成的子群,这是平面一般线性群 $\mathrm{GL}(2,\mathbf{R})$,其元素可由表格

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

表示为简略形式(这也是高斯已经用过的);后来这种表称为矩阵.关系(30)写为向量 $\begin{pmatrix}x\\ y\end{pmatrix}$ 与 $\begin{pmatrix}X\\ Y\end{pmatrix}$(§1,B)之间的一个单个方程:

$$\begin{pmatrix}X\\ Y\end{pmatrix}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+\begin{pmatrix}a\\ b\end{pmatrix};\tag{31}$$

[137]

并规定如果

$$S=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix},T=\begin{pmatrix}\alpha' & \beta'\\ \gamma' & \delta'\end{pmatrix},$$

则有①

$$ST=\begin{pmatrix}\alpha\alpha'+\beta\gamma' & \alpha\beta'+\beta\delta'\\ \gamma\alpha'+\delta\gamma' & \gamma\beta'+\delta\delta'\end{pmatrix}.\tag{32}$$

必须注意,为使 $A(2)$ 的子集 Γ 是一个子群,只是它含有恒等变换

$$I=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

以及 Γ 中两个变换 S,T 之积也在 Γ 中是不够的.例如,如果

① 为与两个函数的复合 $f\cdot g$(§1,C)保持记号一致,这里 ST 意味着先作变换 T,再作变换 S.

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

则对每个自然数 k 有

$$S^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

从而任一乘积 $S^h S^k = S^{h+k}$在 Γ 中,但这些变换连同恒等变换 I 不构成一个群,因为 S 的逆

$$S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

不含于该集合中.(这个集合只构成通称的"半群".)

C. "抽　象"　群

我们看到(§1,G),高斯在研究具有整系数和给定判别式的二元二次型类的合成律 $C_1 + C_2$ 时,证明这个合成律是交换的和结合的,并具有中性元 E,而且每个类 C 具有"相反"类 C',满足 $C + C' = E$.现今很显然,我们说这些类构成一个交换群.对于"模 m 的同余类"(§1,F)关于加法

$$\bar{r} + \bar{s} = \overline{r + s},$$

情形也是如此.这时 $\bar{0}$ 是中性元而 $\overline{-r}$ 是 $\bar{r}$ 的"相反"类.这个群的 m 个元素是

$$\bar{1}, \bar{2} = 2\cdot\bar{1}, \bar{3} = 3\cdot\bar{1}, \cdots, \overline{m-1} = (m-1)\cdot\bar{1}, \bar{m} = \bar{0} = m\cdot\bar{1}.$$

这个群是循环的并由 $\bar{1}$ 生成.然而在变换群概念开始流行后很久,看来在很长时间中没有人注意到我们所说的相似性.更有甚者,看来在 1850 年以前,甚至没有人写过实数或复数形成一个加法群,而正有理数形成一个乘法群!这里我们触及到 19 世纪 [138] 数学家在超越数学划分的传统概念方面的困难,这种传统把数学划分为由所研究的数学对象刻画的若干部分:自然数对于算术,方程对于代数,空间和图形对于几何,函数对于分析.为打破这种生硬僵化的划分并达到考虑的是对象之间的关系这种现代

概念,花了 19 世纪整整一个世纪.

在群论情形中,具有整系数的二元二次型理论属于算术,而在很长一段时间内研究几何学的数学家对当代算术学家的工作从不知晓.甚至对数学具有百科全书式知识的克罗内克和凯莱,也未能完全搞出群的一般定义.直到 1882 年才给出有限群或(更一般地)具有有限个生成元[①]的群的定义,而完全一般的群的定义,则直到 1893 年才给出(见 § 3,C,I).

D. 四元数与代数

我们已经看到,群这个概念是在研究各种各样问题的过程中自动地引进的,它以自然的方式显示出所隐藏的基本性质.同样,一旦深入探究代数方程解的愿望出现,复数的引进就是不可避免的.可以说,这两种理论是被激发的.

哈密顿于 1843 年发表的四元数理论则不同,这是数学史上的第一个例子,也是一种理论的原型:新对象的引进不是由于它出现的那个时代的需要,把它引到这个世界上来仅仅出于好奇心,只是"为了了解".

复数演算(§ 1,A)已同化于平面上点的一种演算体系.当这个观念一出现,就有一些数学家提出这样的问题:关于空间中的点有没有类似的演算体系? 直到 1840 年,还没有回答这个问题的认真尝试[②].这时,一直沉思于复数(把它看作实数偶$\begin{pmatrix} a \\ b \end{pmatrix}$)

---

① 元素 $a_1, a_2, \cdots, a_p$ 组成群 G 的"生成元系",如果 G 的每个元素是一些元素 $x_1, x_2, \cdots, x_k$(k 是任意自然数)之积 $x_1x_2\cdots x_k$,其中 x_j 等于某个 a_i 或其逆 a_i^{-1}(表达方式可能有多种).我们也说 $a_1, a_2, \cdots, a_p$ 生成 G.

② 高斯短短的一句谜样的话或许意味着他不相信这类推广是可能的.

[139] 演算的哈密顿,对三元组

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

就上述问题发起有系统的进攻.由于复数之和定义为坐标相加,哈密顿自然把加法的定义取作与向量加法相同:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \\ c + c' \end{pmatrix}. \quad (33)$$

然而问题在于如何定义乘法

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix}.$$

哈密顿自然要求对于实数或复数乘法成立的代数性质[①]仍然得到满足,这就是结合性、交换性和乘法关于加法的分配性.把复数取作他的模型,哈密顿写

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \cdot 1 + b \cdot i + c \cdot j,$$

其中他令

$$1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, i = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, j = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

他把 $1, i, j$ 想象为位于空间中互相垂直的轴上的 3 个单位向量,正如同 $1, i$ 在复数的几何表示中用作平面上两个互相垂直

① 当不涉及 **R** 的序结构时加法和乘法的性质对 **R** 及其"扩充"**C** 都相同这个事实,当时的英国数学家认为这是必然的,他们把它称为"永恒性原理".

的轴上的单位向量.

指引他全部努力的基本考虑是每个复数 $z = a + bi$ 都有向量$\begin{pmatrix} a \\ b \end{pmatrix}$的长度$\sqrt{a^2+b^2}$与之相联系,此长度记作$|z|$(也称为 z 的模).对于复数的乘法我们有"模律"

$$|zz'|^2 = |z|^2|z'|^2, \tag{34}$$ [140]

事实上它只不过是波罗摩笈多恒等式(27)的特殊情形:

$$(a^2+b^2)(x^2+y^2) = (ax-by)^2 + (ay+bx)^2. \tag{35}$$

哈密顿关于类似于复数演算的想法导致他把长度$\sqrt{a^2+b^2+c^2}$同"三元组"$a+bi+cj$ 联系起来,他进而尝试能否找到一种乘法

$$(a+bi+cj)(x+yi+zj) = A + Bi + Cj,$$

使之除了满足上面提到的代数性质外,还满足"模律"

$$(a^2+b^2+c^2)(x^2+y^2+z^2) = A^2 + B^2 + C^2. \tag{36}$$

对于 $y = z = 0$,为满足(36),取 $A = ax$, $B = bx$, $C = cx$ 即可;因此哈密顿认为取

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

为他设计的乘法的中性元是合适的.

另一方面,如果取 $c = z = 0$,则由(35),通过取 $A = ax - by$, $B = ay + bx$, $C = 0$ 即可使关系式(36)成立.于是哈密顿认为,对于他设计的乘法,必须有 $i^2 = -1$;同样通过取 $b = y = 0$,我们看到必须有 $j^2 = -1$.

采纳他所设置的这些条件,结果为:

$$\begin{aligned}&(a+bi+cj)(x+yi+zj)\\&= (ax-by-cz) + (ay+bx)i + (az+cx)j + (bz+cy)ij.\end{aligned} \tag{37}$$

那么,应当怎样设置 ij 才能使右边也是一个三元组$\alpha + \beta i + \gamma j$?

在这一点上，追随一位创造性数学家的试验足迹（对此他曾详细叙述过），是十分有教益的；它比庞加莱在[9]中关于他发现“富克斯”函数的描绘要易于理解得多．哈密顿相继试验取 $ij=1$，$ij=-1$ 甚至走极端取 $ij=0$．但是没有一种情形能使从(37)导出的 A,B,C 的值满足(36)，因为我们有

$$(a^2+b^2+c^2)(x^2+y^2+z^2)$$
$$=(ax-by-cz)^2+(ay+bx)^2+(az+cx)^2+(bz-cy)^2. \tag{38}$$

然而，把这个公式同(37)相比较，他注意到如果在(37)左边（按分配律）展开过程中取 $ji=-ij$，未知元 $k=ij$ 的系数就会等于该多项式的最后一项 $bz-cy$．然后他突然爆出一个直觉，这个问[141]题可能得解，如果假定：

1）我们使之相乘的是四元组

$$a+bi+cj+dk$$

而不是三元组，并连同补充法则 $k^2=-1$；

2）放弃乘法交换性，取

$$ij=-ji=k,$$

保留结合性以及关于加法的（左和右）分配性．由结合性他推出另外的法则：

$$ik=i(ij)=(i^2)j=-j;$$
$$ki=(ij)i=-(ji)i=-j(i^2)=j;$$
$$jk=j(ij)=-j(ji)=-(j^2)i=i;$$
$$kj=(ij)j=i(j^2)=-i.$$

基于这些法则，他有效地得到关于他称之为四元数 $a+bi+cj+dk$ 的“模律”．可以验证

$$(a+bi+cj+dk)(x+yi+zj+tk)$$
$$=(ax-by-cz-dt)+(ay+bx+ct-dz)i+ \tag{39}$$
$$(az-bt+cx+dy)i+(at+bz-cy+dx)k,$$

而我们有恒等式

$$(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+t^2)$$
$$=(ax-by-cz-dt)^2+(ay+bx+ct-dz)^2$$
$$+(az-bt+cx+dy)^2+(at+bz-cy+dx)^2, \quad (40)$$

这是欧拉已经发现过的①.

在辉煌地解决了他所提的问题之后,哈密顿提出他的四元数"会有什么用途"的问题. 在他生命的最后 20 年,在一些或许有点过分热心的门徒的帮助下,他献身于这一探索. 必须承认,他们的结果有些欠缺;他们坚信四元数理论是解决各种问题的万应灵丹,但他们未能成功地向当时的数学界传播他们的信念. 只是在很晚以后,到 19 世纪末叶,四元数才在诸如群的线性表示或李群的结构这些理论中自然地显现出来,而这些理论在哈密顿时代并不存在. [142]

哈密顿的工作的另一后果是寻求"三元组"或"四元组",更一般地,"n 维向量"(§1,B)之间所有可能的"乘法",这种可能性是基于我们不再要求交换性,甚至不再要求结合性,但保留关于向量加法的(右或左)分配性. 这种研究的主要价值在于,同倾向于相信代数演算的通常法则是不可触动的经典态度相反,它表明了在新的"抽象"数学对象上能够进行的"演算"的多样性. 不过探讨这些相当"无缘无故的"代数研究与其他数学领域之间的关系,或者探讨它们对自然科学的应用,几乎都是 1890 年后才开始的. 我们将在 §5,A,Ⅵ中回过来讨论这一点.

---

① 哈密顿并不知道欧拉的这项工作. 他也不知悉勒让德的下述考察:对于(36),如果 A,B,C 是 a,b,c 和 x,y,z 的有理系数线性函数,那么这样的恒等式不可能存在:事实上,如果取 $a=b=c=1,x=4,y=2,z=1$,则数 $63=3\times21$ 应当是 3 个有理数平方之和. 但 63 形如 $8n+7$,而从丢番图时代以来即已知道形如 $8n+7$ 的数不可能是 3 个有理数平方之和.

3. 集合语言与一般结构

A. 集 合 概 念

把具有相同性质的对象搜集为一个"总合"无疑像语言那样古老;对于一个"整体"(它可以是一些对象的一个总合或诸如平面或空间那样的"广延")的某个"部分"的概念,我们也可以这样说.理解和运用这些概念的需要在古希腊数学刚刚发端时就已变得很明显,在欧几里得著名的"共识"中有"整体大于部分"的表述,他在《几何原本》里为比较两段长度、两个角度或两块面积不断地诉诸这一表述.

再者,各个时代的数学家不可能避免谈论他们所讨论的对象构成的集合及其子集,所用名称多种多样:平面上满足某一条件的点的"几何轨迹",欧拉和高斯论著里数或二次型的"类",柯西和伽罗瓦使用的"子群",黎曼论著里任何维数的"流形"(Mannigfultigkeiten),凯莱论著里符号的"集合".德国数学家也用"Gesamtheit"("全体"),"Inbegriff"("总括"),"System"("系统")和"Gebiet"("领域")."Menge"("集合")这个词虽然波尔查诺已经用过,但只是随着康托尔的工作才开始占据主导地位.

然而,当现代读者发现古希腊数学家从不谈论"自然数的集合",也从不谈论"平面上点的集合"(尽管对这些概念我们似乎很熟悉),或许会感到惊奇.理由在于这些集合是"无穷的",这是一个难以刻画的纯粹消极的概念,而且是哲学学派之间无穷无[143]尽争辩的主题(参见第六章,§3).毫无疑问,必须把这个事实看作从欧几里得到柯西的数学家不愿谈论集合概念的理由之一,因为他们害怕卷入他们觉得会是毫无结果的论战之中.当他们面临对象的一个无穷集合时,他们或者像欧几里得关于素数所做的那样,说了在列举有限个以后仍然还有一些就打住,或者像高斯对二元二次型所做的那样,什么也不说,相信他们的读者的

才智.看来波尔查诺是第一位自由谈论无穷集合的数学家;不过直到戴德金才揭开给出这些对象的实在的数学定义这一幕.这一点将在第六章§3中详细讨论.

B. 集 合 语 言

正是戴德金[①],明确地以定义无穷集和自然数为目的,在其著作《数是什么? 数应当是什么?》(该书出版于1888年,但于1878年即已起草)中,在类似"结果综述"的开头部分,引进了非常精确的语言,这与当时模棱两可的习惯大相径庭.虽然这部论著并未引起它应当具有的直接影响,但到20世纪初,数学家们开始逐渐感受到在数学各个领域中采用相似的、统一的语言的需要.随着戴德金之后一些人物的参与,这种语言成为——我们可以这样称呼——朴素集合语言,它在我们的时代被普遍使用,成为数学家们离开它就做不成事的必备工具,除非他们愿意付出冗长累赘语文的代价.

戴德金并不觉得需要以公理形式呈现这种语言;可以看出,对于他,就像对于他的同时代数学家一样,他所列举的基本定义和结果表达了"常识真理".可以说,尽管后来发生了我们将于第六章中谈及的争执,当今数学家仍在以同样的方式使用这种语言.对他们重要的是,它能使他们没有歧义地表达自己的想法.

戴德金并没有给出集合或该集合的元素的真正定义,在这方面他所做的一点也不比欧几里得关于点和线所做的多.基本的关系是一个元素 x 属于一个集合 E,现今我们把它记作 $x\in E$(其否定记作 $x\notin E$),并陈述为"x 属于 E"或"x 是 E 的一个元素".要定义的第一个概念是子集(或部分):A 是 E 的子集,现今写为 $A\subset E$(其否定写为 $A\not\subset E$),如果 A 的每个元素也是 E [144]

① 这里也提出弗雷格的贡献才算公正,他的工作始于同一时期,目的在于以同样方式引进十分精确的集合语言.

的元素.戴德金承认,如果我们同时有 $A\subset B$ 和 $B\subset A$,则有 $A=B$:一个集合完全由其元素确定.他很快越过并与交运算[①];很早就已知道特殊情形下的这两种运算,而布尔也已联系形式逻辑把它编为规程[②](第六章,§5,A).他简单地强调这些运算可用于一个集合的任意的子集族(他清晰地提示,E 的子集是另一个我们现在记作 $\mathfrak{P}(E)$ 的集合的元素).

对于整个数学而言新的、本质的东西是完全一般的*函数*(现今我们喜欢把它称为*映射*)概念.戴德金不是像先前的构思那样把自己局限于一个或几个实变量的实(或复)函数,而是一下走到推广的最大范围:给定任意的两个集合 E,F,E 到 F 中的映射 f 是使 E 中每个元素 x 对应到 F 中一个*一意确定的*[③]元素的一个*法则*,后一元素就是此映射在 x 处的*值*,一般记作 $f(x)$(对特殊的映射也用另外的记号).现在我们有这样的习惯:以 $x\longmapsto f(x)$ 表示映射 f,它常可避免引入新的字母,例如写 $x\longmapsto x^2$,$x\in\mathbf{R}$.现代数学也常用记号 $f:E\to F$ 或 $E\xrightarrow{f}F$,以明确地标志函数 f 定义于其上的集合 E 和 f 取值于其中的集合 F[④].

[145]

戴德金的阐述仍然缺少一个常用的概念,即任意两个集合 E,F 的积(或"笛卡儿积")$E\times F$;这是稍后由康托尔引进的.这是对 E 的所有元素 x 和 F 的所有元素 y 构作的偶 (x,y)[⑤]组成

---

① 不像布尔,戴德金没有谈到集合 E 的子集 A 的补集,即不属于 A 的 $x\in E$ 所成的集合,常记作 $E\setminus A$.他也没有谈到空集,它的引入对于论证是有用的(莱布尼茨已想到这一点).空集现记作 $\varnothing$,并可写 $E\setminus E=\varnothing$.

② 戴德金并未提到布尔,尽管他必定会知道布尔的工作.

③ 戴德金就这样追随黎曼清除了"多值函数"(即对自变量的同一个值能取几个值的函数)这个混乱概念,它是随同一个复数的平方根或对数的定义引进到单复变量解析函数论中的.

④ 因此这个概念包括由狄利克雷定义的最一般的实值函数,柯西的"置换"和所有几何"变换".

⑤ 当然,如果 $E=F$,则当 $x\neq y$ 时必须区别 (x,y) 与 (y,x).

的集合，它是笛卡儿坐标系的自然的、必不可少的推广.我们可以把映射 $f:E\to F$ 这个概念同积 $E\times F$ 的一个子集联系起来：f 的图象 Γ_f 是 $E\times F$ 的由偶 $(x,f(x))$（对 E 的所有成员 x）组成的子集，这是实变量实值函数的经典"图象"（第三章，§8）的显然推广.

对任何自然数 n，以同样的方式定义 n 个集合的积 $E_1\times E_2\times\cdots\times E_n$. n 个都等于 E 的集合之积记作 E^n.

这种语言的重要性在于，它能使数学家们从19世纪后期起前进到其性质完全不确定的对象之间的关系，这些对象只是简单地作为公理式理论本原对象的一些集合的元素.一个结构由遵从一个公理系统的一定数目的这种"本原"关系确定；这样一种结构的理论则循序渐进地显示一些性质，这些性质纯粹是公理系统的结果，并不依赖于满足这些公理的数学对象的本性.这些性质最后形成的理论，其范围可与经典理论媲美.下面我们要给出一些例子.

C. 代 数 结 构

从当代数学家用得最多而且对自然科学和技术问题也有许多应用的结构中选择，我们从称为代数结构的一些例子开始.

（I）群

最古老、最简单的结构是群.前面（§2，B 和 C）已给出某些群的例子，它们是在群的一般概念形成前驱使数学家们研究的.

如同希尔伯特在几何学情形所做的那样处理（第三章，§4），一个群由3个"本原"对象构成：一个集合 G，一个元素 $e\in G$ 以及 G 中的一个合成律即一个映射 $m:G\times G\to G$.关于群有3条公理： [146]

1）合成律 m 是结合的，这就是说，对 $x,y,z\in G$，有

$$m(m(x,y),z)=m(x,m(y,z)).\qquad(41)$$

2）e 是对于 m 的中性元，这就是说，对 $x\in G$，有

$$m(e,x)=m(x,e)=x. \tag{42}$$

3) 每个元素 $x \in G$ 有逆元 x',换言之有

$$m(x,x')=m(x',x)=e. \tag{43}$$

一个群称为交换的,如果加之还有,对 $x,y \in G$,

$$m(x,y)=m(y,x). \tag{44}$$

如果 G 只有有限个元素,则称它为有限群;其元素个数称为 G 的阶.

群 G 的子集 H 称为 G 的一个子群,如果对 $x,y \in H$ 有 $m(x,y) \in H$,$x' \in H$(即积与逆在 H 中).显见所有这些都是关于置换群已知规律(§2,B)的推广.用同样的论证可以证明,如果 G 是有限群,H 是 G 的子群,则 H 的阶整除 G 的阶.

通常我们写 xy 以代替 $m(x,y)$,用 x^{-1} 代替 x'.我们逐次定义 $x^2=xx$,$x^3=x^2x$,$\cdots$.如果存在最小的自然数 k,使得 $x^k=e$,则元素 $e,x,x^2,\cdots,x^{k-1}$ 互不相同并构成一个 k 阶交换子群,它是循环的.我们也称 k 是 x 的阶.如果 G 是有限群,则 G 的所有元素是有限阶的,且它们的阶整除 G 的阶.这样,如果 n 是 G 的阶,则对每个 $x \in G$ 有 $x^n=e$.

对于 G 为交换的情形,我们常写 $x+y$ 来代替 $m(x,y)$,以 0 代替 e,以 $-x$ 代替 x';于是定义

$$2x=x+x, 3x=2x+x, \cdots.$$

当群的合成律写为 $x+y$ 时,常称该群为加法群,而当合成律写为 xy 时,则称为乘法群.

如果 G,H 是两个群,则对笛卡儿积 $G \times H$,通过令

$$(x,y)(x',y')=(xx',yy'),$$

可定义一个结构.我们称这样定义的群为 G 与 H 的积群.这个概念可直接推广到任何个数的群.

(Ⅱ) 环

我们已注意到,在整数的集合(记作 **Z**)、实数的集合(记作

R)或复数的集合(记作 C)中,有加法与乘法两种合成律.数论中的问题导致欧拉,后来导致高斯以更系统的方式考虑 C 的有别于 R 或 Z 的子集,对于它们加法与乘法两者都不越出所给子集.这种子集称为环(戴德金把它们称为"目").环的最早例子由形如 $a+b\sqrt{D}$的数构成,这里 D 是一个固定的不等于一整数平方的整数,a,b 是变动的整数(如果 $D<0$,则 $a+b\sqrt{D}=a+ib\sqrt{-D}$是复数).事实上我们有 [147]

$$(a+b\sqrt{D})+(a'+b'\sqrt{D})=(a+a')+(b+b')\sqrt{D}, \quad (45)$$

$$(a+b\sqrt{D})(a'+b'\sqrt{D})=(aa'+bb'D)+(ab'+ba')\sqrt{D}. \quad (46)$$

这样的数构成的环记作 $\mathbf{Z}[\sqrt{D}]$.

高斯详细考察了数 $a+bi(=a+b\sqrt{-1})$的环的结构,现在称这样的数为高斯整数.他表明,在这个环中有非常接近于通常算术的"算术":可以定义"素数",并且每个高斯整数有唯一的方式分解为素因数(附录 3,B).正是这种因数分解"解释"了费马和欧拉关于 $x^2+y^2=n$ 的整数解的定理(第四章,§2,A).但是进一步的研究表明这些算术性质对许多 D 值不再成立(例如 D=-5).经过半个世纪的努力才克服了这个困难;而且正如"用根式"解代数方程问题是群论的起源(§1,D)一样,关于环 $\mathbf{Z}[\sqrt{D}]$和由"单位根"(满足方程 $x^n-1=0$ 的复数)生成的环的研究是代数数论这个伟大的现代理论的起源.

上面所举并不是能定义两种合成律的仅有的例子.定义在任意集合 E 上的实值函数的集合关于§1,C 中定义的加法 $f+g$ 与乘法 fg 构或一个环,任何个数的实或复变量的多项式也构成一个环.

这些考察在 20 世纪初叶导致交换环这种一般结构的定义.这里"本原"对象是一个集合 A 与两种"合成律",分别记作 $(x,y)\longmapsto x+y$(加法)与$(x,y)\longmapsto xy$(乘法),其公理为:仅

仅关于加法，A 是加法群，而乘法是结合的、交换的且关于加法是分配的.

(Ⅲ) 域

[148] 某些环(但非所有环①)具有单位元 e，使对每个 x，有 $ex=x$. 而从 19 世纪初期起，阿贝尔和伽罗瓦发掘出这样的复数集(但未给出名称)K：它是包含有理数集 $\mathbf{Q}$ 的环，且对每个不等于零的 $x\in K$②，其逆(倒数)$1/x$ 也在 K 中. 从戴德金时代以来，这些集合被称为交换域或 $\mathbf{C}$ 的子域. 集合 $\mathbf{Q}$，$\mathbf{R}$ 和 $\mathbf{C}$ 都是域. 阿贝尔和伽罗瓦讨论的域 K 通过取有限个任意的复数 $a_1, a_2, \cdots, a_m$ 得到：K 是数 $R(a_1, a_2, \cdots, a_m)$ 的集合，其中 $R=P/Q$ 是两个 m 元任意次有理系数多项式 P，Q 之商，并满足条件 $Q(a_1, a_2, \cdots, a_m)\neq 0$. 事实上，对于满足 $Q(a_1, \cdots, a_m)\neq 0$，$Q_1(a_1, \cdots, a_m)\neq 0$ 的多项式 P, Q, P_1, Q_1，我们有

$$\frac{P}{Q}+\frac{P_1}{Q_1}=\frac{PQ_1+P_1Q}{QQ_1},\ \frac{P}{Q}\cdot\frac{P_1}{Q_1}=\frac{PP_1}{QQ_1},$$

而乘积 $Q(a_1, \cdots, a_m)Q_1(a_1, \cdots, a_m)\neq 0$. 如果加之还有 $P(a_1, \cdots, a_m)\neq 0$，则数

$$\frac{P(a_1, \cdots, a_m)}{Q(a_1, \cdots, a_m)}$$

具有逆

$$\frac{Q(a_1, \cdots, a_m)}{P(a_1, \cdots, a_m)}.$$

---

① 一个例子是 $\mathbf{Z}$ 中由偶数构成的子环 $2\mathbf{Z}$. 两个偶数之积是偶数，但不存在偶数 e，使得 $e(2n)=2n$.

② 环的公理蕴涵对一切 $x\in A$，$0x=0$，因为对所有 $x\in A$，我们有

$$x\cdot x=(x+0)\cdot x=x\cdot x+0\cdot x,$$

而由于 A 关于加法是群，因此由上式得 $0\cdot x=0$.

我们写 $K=\mathbf{Q}(a_1,\cdots,a_m)$,并称 K 为由 $a_1,\cdots,a_m$ 生成的域.此外,在这一定义中,自始至终域 $\mathbf{Q}$ 可代之以复数域 $\mathbf{C}$ 的任一子域.

还应注意到,当 a,b 是有理数时,数 $a+b\sqrt{D}$ 也构成一个域,因为不能有 $a+b\sqrt{D}=0$,即 $a^2-b^2D=0$,除非 $a=b=0$,这是由于 D 不是一个平方.这样,由于 $a+b\sqrt{D}\neq 0$,我们能写

$$\frac{1}{a+b\sqrt{D}}=\frac{a}{a^2-b^2D}-\frac{b}{a^2-b^2D}\sqrt{D},$$

它也有 $c+d\sqrt{D}$ 的形状,其中 c,d 为有理数.

(Ⅳ) 非交换环和非交换域

哈密顿四元数(§2,D)的集合 $\mathbf{H}$"近于"一个域:事实上,如果四元数

$$q=a+bi+cj+dk$$

不等于零,则有 $a^2+b^2+c^2+d^2\neq 0$.于是关于乘法的公式(39) [149] 表明四元数

$$q'=\frac{a-bi-cj-dk}{a^2+b^2+c^2+d^2}$$

满足 $qq'=q'q=1$.换言之,每个非零四元数具有(右或左)逆.

这就导致推广交换环的定义:一个(交换或非交换的)环是一个集合 A,并在 A 中定义了:(i) 一个加法,而 A 关于此加法是一交换群;(ii) 一个结合乘法,它不一定是交换的,但关于加法是右和左分配的,就是说,有

$$x(y+z)=xy+xz,(y+z)x=yx+zx. \tag{47}$$

一个(不一定交换的)域是一个环 K,其乘法具有单元位 e,且每个 $x\neq 0$ 具有右和左逆元 x',换言之,具有满足 $xx'=x'x=e$ 的 x'.通常我们写 1 来代替 e,并以 x^{-1} 代替 x'.注意由这些条件推出 K 的非零元的集合 K^* 关于乘法是一个群.四元数集 $\mathbf{H}$ 提供了非交换域的例子.

D. 序 结 构

我们已看到(第三章, §4),直线上点的顺序这一想法(或“位于两个点之间”的点这一概念)是基本几何概念之一.这是比较“大小”或比较整数这种直观概念的数学“模型”.在19世纪最后30年中,康托尔和戴德金从各个方向加深和推广了这个概念.

康托尔在其数学生涯的开始阶段,追随黎曼研究分析中这样的问题,其结果可能先验地依赖于“例外”实数的一个集合 E(附录5,D).他首先证明,如果 E 是有限的,则这些数对结果没有影响.但当 E 是无穷集时,只有作出一些补充假定,他才能达到相同的结论;这导致他认识到,如果对实数按其元素的序结构来考虑,那么由实数组成的各种集合有着极大的多样性.例如,如果 E 是有理数的集合,那么在 E 的两个数之间总有另一个数(甚至无穷多个另外的数).另一方面,考虑数 $1/n$ $(n=2,3,\cdots)$ 组成的集合 E,它就没有上述性质.注意此集合有最大数 $1/2$,[150] 但没有最小数;可是数 $1-1/n$ $(n=2,3,\cdots)$ 组成的集合 F 则恰好相反.并 $E\cup F$ 既没有最大数,也没有最小数.这样一些例子激励康托尔对序观念进行更一般的研究是可以理解的.

就戴德金而言,他对数论的探讨导致研究两个自然数 a 与 b 之间“a 整除 b”的关系.如果把 a 整除 b 写为简略形式 $R(a,b)$,则我们有下列三条性质:

1) $R(a,a)$ 恒为真(这称为关系 $R(a,b)$ 的自反性);

2) 如果有 $R(a,b)$ 和 $R(b,c)$,则有 $R(a,c)$(关系 $R(a,b)$ 的传递性);

3) 如果同时有 $R(a,b)$ 和 $R(b,a)$,则必有 $a=b$(这个性质有时称为反称性).

注意实数之间通常的关系 $a\leqslant b$(它曾被康托尔研究过)显然满足这三条性质.然而,虽然我们总有或者 $a\leqslant b$ 或者 $b\leqslant a$

成立,但对关系"a 整除 b"情形却不是如此:3 不整除 7 且 7 不整除 3.

这些例子以及其他许多例子导致任一集合 E 的元素之间序(或偏序)关系这个一般概念:它是所给集合两个元素之间满足上述性质 1),2),3)的一个关系 $R(a,b)$;这三条性质就是序结构的公理.如果加之 $R(a,b)$还满足下述公理:

4) 对 E 的每对元素 a,b,或者有 $R(a,b)$,或者有 $R(b,a)$(如果两者均成立,则由 3)有 $a=b$),则称 $R(a,b)$是一个全序关系.

与一般人可能设想的相反,在数学中我们遇到(偏)序关系远多于遇到全序关系.最常遇见的是"部分"与"整体"之间的包含关系 $A\subset B$;欧几里得经常用到它,它满足上述 1),2),3),但只要"整体"含有多于两个元素,那么它就不满足 4):如果 D_1,D_2 是平面上两条不同的直线,则两者都不是另一方的子集.

E. 度量空间与拓扑概念

我们已看到(第三章,§9),数列的极限和实变量实值函数的连续性概念,在很长时间里处于模糊状态,最后在 19 世纪初得到精确表述.这是两个现称为"拓扑概念"的概念.

另外的拓扑概念在 1860 年后出现于魏尔斯特拉斯和康托尔的论著中(也出现于戴德金所作但生前未发表的笔记中).魏 [151] 尔斯特拉斯在讨论分析基础时,在其讲授中定义实数 x_0 的邻域:这是实数集的一个子集,它包含未特别指定但长度不等于零的区间$[x_0-\alpha,x_0+\alpha]$.戴德金在其笔记中由此导出开集(在那个阶段他称之为"区域")概念:这是一个子集 U,它是自身中每个点的邻域.戴德金甚至通过以圆盘或球代替区间把这个定义推广到平面上和空间中.就康托尔而言,关于"例外"集的研究(上述 C)导致他考察集合 $\mathbf{R}$ 的任意子集 E 的性质,他定义 E 的导集 E'为满足下述条件的点 x_0 的集合:x_0 的每个邻域含有 E

的无穷多个点(因此,如果 E 是有限集,则 E' 是空集).然后他取 E' 的导集 $(E')' = E''$,如此继续,并审察可能发生的各种情形.例如,对于数 $1/n(n=2,3,\cdots)$ 组成的集合 E,E' 退化为只含有一个点 0;对于形如 $1/n$ 和 $1-1/n$ 的数组成的集合 E,E' 含有两个点 0,1;在这两种情形,E'' 都是空集.另一方面,如果 E 是介于 0,1 之间的有理数的集合,则 E' 是整个区间 $[0,1]$,随后的导集都等于这个区间.

然而,分析尤其是变分法①的需要使得在相当早之前,这些"拓扑"概念就被推广到与实数不同的别的对象:我们需要对向量、实值函数、曲线、曲面等等谈论"极限"或"邻域".是弗雷歇于1906年发现了一个非常简单而又完全一般的结构,它使给出分析的绝大多数问题中自然产生的"拓扑"概念的意义成为可能,这就是现称为度量空间的结构.这种结构的"本原对象"是一个任意的集合 E 和从 $E\times E$ 到 $\mathbf{R}$ 中的一个映射 $(x,y)\longmapsto d(x,y)$,它满足下列三条公理:

Ⅰ) $d(x,x)=0$;如果 $x\neq y$,那么 $d(x,y)>0$;

Ⅱ) $d(y,x)=d(x,y)$;

[152] Ⅲ) $d(x,z)\leqslant d(x,y)+d(y,z)$.

很清楚,在欧几里得空间中,两点 x,y 之间的距离满足这些公理,因为可以辨认出,Ⅲ)表达了三角形边之间熟知的不等式(欧几里得,《几何原本》,第Ⅰ卷命题 20).这样,度量空间的元素通常称为点,而数 $d(x,y)$ 称为距离.

值得注意的是,所有"拓扑"概念能仅用上面的公理来定义.例如,E 中的无穷点列 (x_n) 以点 y 为其极限(或"趋于"y),如果

① 这门学科诞生于17世纪末,其目的是求解与变动曲线或曲面有关的数的极小值或极大值的计算这个困难问题;例如,在平面上具有给定长度的不自交闭曲线中求出包围面积最大的曲线,或在曲面上,诸如椭球面上,求连接两个点的长度为极小的曲线.

数列$(d(y,x_n))$趋于零.设E,F是两个度量空间,映射$f: E \to F$是连续,如果每当E中点列(x_n)趋于极限y,总有F中点列$(f(x_n))$趋于$f(y)$.点$x_0 \in E$的一个邻域是一个集合V,使对某个充分小的数$\alpha > 0$,V包含满足$d(x,x_0) \leqslant \alpha$的元素$x \in E$构成的集合;等等.

要注意,同一集合上两个不同的距离可能产生相同的拓扑概念.例如,在平面$\mathbf{R}^2$上,对两个点(x_1,y_1),(x_2,y_2),我们能取欧几里得距离

$$\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$$

为两者之间的距离,也能取数

$$|x_1-x_2|+|y_1-y_2|$$

为两者之间的距离.我们称集合E上的两个距离$d_1(x,y)$与$d_2(x,y)$是等价的,如果对于这两个距离的邻域概念相同.E上等价距离的集合组成称为可度量化空间的结构.

度量空间具有极大的多样性,此处我们不能描述它在分析、微分几何、代数拓扑或甚至数论中所起的作用.我们在这里只提及除空间$\mathbf{R}^n$外或许用得最多的空间,这就是希尔伯特空间,它是量子力学的数学基础(附录4,C).

F. 结构的叠置和分离

前面几节描述的是最简单的结构,可以称为"纯粹"结构.然而,19世纪进程中出现了一些结构类型,它对解决经典问题甚至更加有用.它们可以称为"混合"结构,其中可以有几个赋予结构的"本原"集合,也可以在同一个集合上有几个结构.自然有把这些结构互相联系起来的公理.这种"混合"结构的数目正在不断增多,我们将只提几个例子.[153]

最早的例子是向量空间结构,它有两个"本原"集合,即一个交换域K(本节CⅢ)和一个使用加法记号的交换群E(本节C

Ⅰ).还有一个本原对象,即 $K\times E$ 到 E 中的一个映射 $(\lambda,x)\longmapsto\lambda x$($\lambda x$ 也记作 $\lambda\cdot x$),它称为标量乘法,遵从下列公理:

$$\lambda(x+y)=\lambda x+\lambda y;\tag{48}$$

$$(\lambda+\mu)x=\lambda x+\mu x;\tag{49}$$

$$\lambda(\mu x)=(\lambda\mu)x;\tag{50}$$

$$1\cdot x=x.\tag{51}$$

典型例子是 n 维向量的集合 $E=\mathbf{R}^n$(§1,B),这里 $K=\mathbf{R}$,标量乘法是一个位似变换(§1,B,E).$\mathbf{R}^n$ 上的向量空间结构由凯莱和格拉斯曼于1843年提出,它是数学各个领域中最广泛、最有用的结构之一.

这个结构可以有多种推广途径.在公理(48)到(51)中,不一定必须假定 K 是域,它可以代之以具有单位元[①]的交换环 A.于是就有 A 模结构,它处于戴德金于1860—1870年创立的代数数论和克罗内克于同一年代构思的代数函数论的中心.

更进一步,通过不仅在 E 上设置加法群结构,而且还设置(交换或非交换)环结构,就可能"丰富"向量空间结构.这时必须在(48)到(51)上增添一个把环 E 中的乘法同标量乘法联系起来的新公理,即

$$\lambda(xy)=(\lambda x)y=x(\lambda y).\tag{52}$$

这时可以说我们定义了一个 A 代数结构(在19世纪它被称为"*超复数系*").哈密顿四元数集(§2,D)是一个 $\mathbf{R}$ 代数.

另一方面,如果我们以使用乘法记号的(不一定是交换的)群 G 代替 K 来"减弱"向量空间结构,自然应取消公理(49),因为它不再有意义;这样得到的结构称为*带算子加法群* E(算子指 G 中的元素).我们甚至可以取消 E 上的所有结构,只保留 [154] $G\times E$ 到 E 中的映射 $(\lambda,x)\longmapsto\lambda\cdot x$,还有公理(50)和(51);于是我们有一个*带算子群 G 的集合* E.这时称 G *作用于*(或*运算*

① 当 A 没有单位元时,当然要取消公理(51).

于)E,映射$(\lambda,x)\longmapsto\lambda\cdot x$ 称为 G 在 E 上的作用.

最后这个结构的重要性来源于下述事实：自从克莱因时代(1873 年)以来,正是这种结构承担为群 G 的几何学,从而统一了 19 世纪初叶以后发展起来的各种几何概念(实射影几何、复射影几何以及非欧几里得几何).对于克莱因,2 维或 3 维古典欧几里得几何是这样一种“几何”,其中 E 是平面 $\mathbf{R}^2$ 或空间 $\mathbf{R}^3$,而 G 是位移群(§2,B).对于克莱因,一条“几何定理”是 E 的元素或子集之间的在 G 的作用下不变的一个关系,即对每个算子 $\lambda\in G$,当该关系涉及的 E 中元素 $x,y,z,\cdots$代之以 $\lambda\cdot x$, $\lambda\cdot y,\lambda\cdot z,\cdots$时,它仍然成立;对于 E 的子集情形也是如此.显然这正是欧几里得几何学所有定理的共性.

作为完全不同的“几何”的例子,让我们引用庞加莱半平面 H,H 是满足 $y>0$ 的复数 $z=x+iy$ 的集合.群 G 是保向平面线性群,记作 $\mathrm{GL}^+(2,\mathbf{R})$.$G$ 由矩阵(§2,B)

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

组成,这些矩阵满足：$\alpha,\beta,\gamma,\delta$ 都是实数,且 $\alpha\delta-\beta\gamma>0$. G 在 H 上的作用由

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\cdot z=\frac{\alpha z+\beta}{\gamma z+\delta}\tag{53}$$

定义.(其结果总是有意义的,因为对任一 $z\in H$ 不可能有 $\gamma z+\delta=0$. 首先,γ,δ 不能都等于零;其次,如果 $\gamma\neq0$,则不能有 $z=-\delta/\gamma$,因为 z 不是实数.)可以直接验证

$$\frac{\alpha(x+iy)+\beta}{\gamma(x+iy)+\delta}$$

的虚部是

$$\frac{(\alpha\delta-\beta\gamma)y}{(\gamma x+\delta)^2+\gamma^2y^2}.$$

由于 $\alpha\delta-\beta\gamma>0$ 且 $y>0$,所以此虚部大于零.我们在讨论非欧

几里得几何时还会遇到这种“几何”(第六章，§1,C). 研究 $GL^+(2,\mathbf{R})$的某些子群在 H 上的作用，是庞加莱的富克斯函数理论(1881)的基础.

作为几个结构“叠置”在同一个集合上的例子，让我们引用*有序交换域*结构. 这是同时在一个集合 K 上考虑交换域结构和 [155] 记作 $x \leqslant y$ 的全序结构. 联系这两个结构的公理是：

$$\text{如果 } x \leqslant y\text{，则对任何 } z \in K \text{ 有 } x + z \leqslant y + z\text{；} \tag{54}$$

$$\text{如果 } x \leqslant y\text{，且 } z \geqslant 0\text{，则 } xz \leqslant yz. \tag{55}$$

这里我们能辨认出实数系关于通常序关系的两条公理(第三章，附录 2)，其中第一个本质上就是欧几里得关于量的“共识”之一. 实数域 $\mathbf{R}$ 的所有子域都是有序域；但还有许多别的有序域(附录 3,E). 也有这样的交换域，例如复数域 $\mathbf{C}$，对于它没有满足公理(54)和(55)的序结构(附录 3,E).

如果 G 是使用加法记号的交换群，且在 G 上定义了一个满足(54)的(全或偏)序，则称在 G 上定义了*有序群*结构. 例如，对 $G = \mathbf{R}^2$，通过取序关系

$$(x_1, x_2) \prec (y_1, y_2)$$

为关系

$$x_1 \leqslant y_1 \text{ 和 } x_2 \leqslant y_2 \tag{56}$$

(这里 $\leqslant$ 为实数之间的通常关系)，定义一个偏序群结构.

然而，通过取关系

$$x_1 < y_1 \text{ 或}(x_1 = y_1 \text{ 且 } x_2 \leqslant y_2) \tag{57}$$

代替(56)，我们也能在 $\mathbf{R}^2$ 上定义一个满足(54)的全序结构. 关系(57)就是称为*字典序*的序关系.

另一个“叠置”结构的例子是*可度量化拓扑群*，它是在“本原”集合 G 上定义一个(交换或非交换)群结构和一个可度量化空间结构，两者之间满足下述公理：

> 如果 G 的元素的序列(x_n)，(y_n)具有极限，则序列$(x_n y_n^{-1})$具有极限，且若 $x = \lim x_n$，$y = \lim y_n$，则 $xy^{-1} =$

$\lim(x_n y_n^{-1})$. (58)

例如，对于前面定义的矩阵群 $GL^+(2,\mathbf{R})$，如果我们对两个矩阵

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

定义两者之间的距离为数

$$|a'-a|+|b'-b|+|c'-c|+|d'-d|,$$

则上述公理得到满足.

第一批拓扑群是S·李于1873年在"连续群"的名称下考虑的；他应用这种群来解微分方程的积分问题. 它在分析、微分几何甚至数论、相对论、量子物理中变得具有极大的重要性. [156]

可以用同样方式通过下述公理定义可度量化拓扑环 A：

如果 A 中的两个序列 (x_n)，(y_n) 具有极限，则序列 (x_n-y_n) 和 (x_ny_n) 都有极限，且若 $x=\lim x_n$，$y=\lim y_n$，则有 $x-y=\lim(x_n-y_n)$，$xy=\lim(x_ny_n)$. (59)

最简单的例子显然是域 $\mathbf{R}$，但这种环出现于数学的所有领域中，甚至出现于数论中：在整数环(在高斯看来，这尤其是"不连续的"对象) $\mathbf{Z}$ 上，存在无穷多个距离(对每个素数有一个)，对于它们0到一个整数 $n\neq0$ 的距离能成为任意小但不等于零(附录4,D). 尽管这看起来可能挺奇怪，但这些距离在数论中具有相当的重要性.

可以看出，特别地在实数集上，各种结构——域结构、序结构、拓扑结构交汇在一起. $\mathbf{R}$ 上还有测度概念，由于它的专门性质，我们不能在此地细述；但因它不但能使区间 $[a,b]$ 对应一个非负的数(简单地就是它的"长度" $b-a$)，而且能使分析的各种主要问题中出现的所有集合对应一个非负数，因此在分析中起着极其重要的作用. 基本上是由于康托尔，我们才能从多样性的观点考察实数集；而在19世纪的大部分时间中，总是用"连续统"这个词而对 $\mathbf{R}$ 上定义的各种结构不加区别. 我们会在第六

章中看到,这引来非同寻常的混乱.

4. 同构与分类

A. 同　　构

有没有可能对两个集合 E,F 上相同类型的结构,例如群结构加以比较?首先必须比较集合 E,F 本身而不管其上的结构.例如,如果 E,F 是有限集,则当它们不具有相同数目的元素时,就把它们看作不同.从康托尔时代以来,基于一一映射概念,这个想法被用于所有集合;一一映射是具有下述性质的映射 f: $E\to F$:对任一元素 $y\in F$,存在唯一的元素 $x\in E$,使得 $y=f(x)$.例如,$x\longmapsto x^3$ 是 $\mathbf{R}$ 到其自身的一个一一映射,但 [157] $x\longmapsto x^2$则不是,因为如果 $y<0$,则它不可能是一个平方,而如果 $y>0$,则存在两个数 $x=\sqrt{y}$,$x=-\sqrt{y}$使得 $x^2=y$.康托尔称两个集合 E 与 F 是等势的,如果存在一个一一映射 f: $E\to F$.此时也有一个一一映射 f^{-1}: $F\to E$,就是使得 $y\in F$ 对应到满足$f(x)=y$的唯一元素 x 的映射.两个有限集等势当且仅当它们具有相同数目的元素,而有限集不可能与无穷集等势.康托尔还发现了下述值得注意的事实:存在不等势的无穷集 E,F(参阅第六章,§3,B).

现在让我们假定 E,F 是两个群.E 与 F 称为同构的,如果存在一个一一映射 f: $E\to F$,使得加之还有,在映射 f 下对应于E 中两个元素 x,y 之积xy 的是积$f(x)f(y)$,换言之,

$$f(xy)=f(x)f(y). \tag{60}$$

这样的一一映射 f 称为 E 到 F 上的一个同构,而其"逆" f^{-1}: $F\to E$ 也是一个同构,因为如果 x',y'是F 的两个元素且 $x=f^{-1}(x')$, $y=f^{-1}(y')$,则有 $x'=f(x)$, $y'=f(y)$,因而由(60)有

$$x'y'=f(x)f(y)=f(xy).$$

由此得到

$$f^{-1}(x'y')=xy=f^{-1}(x')f^{-1}(y').$$

例如,考虑两个相同阶数的 n 阶循环群(§3,CⅠ)E,E'.于是 E(对应地,E')有一元素 a(对应地,b)使得 E(对应地,E')由不同的元素 $e,a,a^2,\cdots,a^{n-1}$(对应地,$e',b,b^2,\cdots,b^{n-1}$)组成,并且 a^n 等于中性元 e(对应地,b^n 等于中性元 e').通过取 $f(e)=e',f(a^k)=b^k(1\leqslant k\leqslant n-1)$定义一一映射 f: $E\to E'$.我们来证明确实有 $f(xy)=f(x)f(y)$.如果 x 或 y 是中性元 e,此式显然成立.否则,存在 $h,k,h\leqslant n-1,k\leqslant n-1$,使得 $x=a^h,y=a^k$.由此 $xy=a^{h+k},f(x)f(y)=b^{h+k}$.如果 $h+k\leqslant n-1$,则由 f 的定义确实有 $f(xy)=f(x)f(y)$;如果 $h+k\geqslant n$,则可写 $h+k=n+j,0\leqslant j\leqslant n-2$,由此

$$a^{h+k}=a^na^j=a^j,b^{h+k}=b^nb^j=b^j,$$

因而

$$f(xy)=f(a^{h+k})=f(a^j)=b^j=b^{h+k}=f(x)f(y).$$

这就证明了下述定理:所有同阶的循环群是同构的.这样,模 n 同余类的加法群(§2,C)与循环置换$(12\cdots n)^k$ 的群(§2,B)这两个表面上不同的群是同构的.

从一个群 E 到另一个群 F 上的同构的存在性的巨大意义在于,对于 E 的元素或子集只借助群公理表达的任何性质,我们可立即推出同样的性质对于 F 的由所给同构对应出的元素或子集也成立,而后者有时却很不显见. [158]

然而同构的存在性并不总是那么显然.以(正、负或零)实数的加法群 $\mathbf{R}$ 与正实数的乘法群 $\mathbf{R}_+^*$ 为例,直到17世纪,才发现 $\mathbf{R}_+^*$ 到 $\mathbf{R}$ 上的同构——底数 $a(a>1)$的对数 $x\longmapsto\log_a x$.事实上这是一一映射,而且不管两个群的写法不同,关系式

$$\log_a(xy)=\log_a x+\log_a y \tag{61}$$

不是别的,正是方程(60).

另一方面,如果考虑加法群 $\mathbf{R}$ 的由正、负有理数和零组成

的子群 $\mathbf{Q}$ 与乘法群 $\mathbf{R}_+^*$ 的由正有理数组成的子群 $\mathbf{Q}_+^*$,那么这两个群却是不同构的.因为对每个 $x \in \mathbf{Q}$,存在元素 $y \in \mathbf{Q}$,使得 $2y = x$.如果存在同构 f: $\mathbf{Q} \to \mathbf{Q}_+^*$,那么对于每个 $x' \in \mathbf{Q}_+^*$,会存在一个且只存在一个 $x \in \mathbf{Q}$,满足 $f(x) = x'$,而如果 $2y = x$ 且 $f(y) = y'$,则按定义就会有

$$y'^2 = f(2y) = f(x) = x'.$$

但从毕达哥拉斯时代起就已知道当 $x' = 2$ 时这是不成立的(参见第三章,§2).

对于§3中考虑的每种结构,能以同样方式定义同构概念.对于两个环 A,B,从 A 到 B 上的一个同构 f 是一个一一映射,使对 A 中任何 x,y,有

$$f(x+y) = f(x) + f(y), f(xy) = f(x)f(y). \qquad (62)$$

例如,设 E,F 是两个等势的集合,A,B 分别是由定义在 E,F 上的实值函数构成的环.令 $\sigma: F \to E$ 是一个一一映射,则使定义于 E 上的函数 $x \longmapsto u(x)$ 对应到定义于 F 上的函数 $y \longmapsto u(\sigma(y))$ 的映射 $u \longmapsto u \circ \sigma$ 是 A 到 B 上的一个同构,此事易于证实.

对于两个有序集 E,F,E 到 F 上关于序结构的同构是一个一一映射 f: $E \to F$,使得关系 $x \leqslant y$ 等价于 $f(x) \leqslant f(y)$①.例如,映射 $x \longmapsto 2x$ 是区间 $[0,1]$ 到区间 $[0,2]$ 上的一个同构. [159]

如果 E,E' 是两个度量空间(§3,E),d,d' 分别是这两个空间中的距离,那么从 E 到 E' 上的度量空间同构(也称为等距同构)是一个一一映射 $f: E \to E'$,使得

$$d'(f(x), f(y)) = d(x, y).$$

① 当并未假定 E,F 上的序结构是全序时,一个一一映射 f: $E \to F$ 可能使得关系 $x \leqslant y$ 蕴涵 $f(x) \leqslant f(y)$,但并不等价于 $f(x) \leqslant f(y)$.例如,如果 E,F 两者都是自然数集,f 是恒等映射,则当在 E 上取关系"x 整除 y"而在 F 上取通常序关系 $x \leqslant y$ 时,显见 x 整除 y 蕴涵 $x \leqslant y$,但反之不成立.

现代分析的伟大进展之一是揭示积分论中定义的某个度量空间同构于一个希尔伯特空间.

如果我们只有 $d''(f(x),f(y))=d(x,y)$,这里 d'' 是等价于 d'(§3,E)的一个距离,则称 f 为可度量化空间 E 到可度量化空间 E' 上的一个同胚.它相当于说 f 和 f^{-1} 都连续.

B. 分 类 问 题

每当定义了一个结构并证明它对求解经典问题有用时,我们自然就会提出对满足同一公理系统 Σ 的所有结构加以分类的问题.当然,两个同构的结构不加区别.换言之,任务在于确定满足 Σ 的结构的一个总集,其中任何两个互不同构,而每个满足 Σ 的别的结构同构于此总集中的某个结构.

这里提出的是一个理想目标,它只在少数情形中得到实现.1865 年,谢林受高斯遗留的论文的启发(当时他正在编集高斯的论著),阐明了具有给定判别式的二次型类(§2,C)的交换群的结构.两年后克罗内克注意到,用同样的方法可对所有有限交换群进行分类.结论非常简单:这些群是有限个阶为素数之幂的循环群之积.

很快认识到对有限非交换群分类显现出程度相当不同的困难.在同构范围内确定给定阶为 N 的所有的群这个问题已经非常艰难:没有人能得到数 N 超过几千时的结果.于是求助于退回到满足附加的有趣性质的某种类型的有限群.例如,阶为素数 p 之幂的群(称为 p 群)已得到很多研究,它们具有值得注意的、非常有限制性的性质;尽管如此,我们离描述所有这种群仍十分遥远.

一个可以追溯到伽罗瓦的问题是确定非交换有限单群,在某种意义上它同 p 群相反;它们组成所谓的“基本构件”,由之出发所有别的有限群从理论上说都能重新构造出来(附录 2, [160] D).寻求确定所有这些群(在同构范围内)持续了 150 年之久.

伽罗瓦发现了最小的单群，它有 60 个元素；但直到 1983 年，才做到完全枚举所有这些群，其证明（正在验证之中[1]）长达一万余页！

当然可以预料，无限群会有复杂得多的多样性，只有设置补充公理我们才能指望对其中某些群加以分类. 最早的成就之一是某种也称为单群（但与有限单群的定义不同）的李群（§3，F）的完全分类. 所有这些群在 1890 至 1930 年间由 S·李、W·基灵、E·嘉当给出描述.

差不多在同一期间，类似地得到了某种也称为单的 **R** 代数和 **C** 代数（§3，F）的分类.

至于域，我们知道如何完全确定的只是有限域，本质上（没有有关的术语）这是由高斯（他在这个论题上从未发表任何东西）和伽罗瓦发现的. 有限域的元素的数目总是一个素数的幂，而对所有素数幂 p^n，在同构范围内只有唯一一个域结构. 这些域必须是交换的. 最近它们在信息论中构造所谓"纠错码"时变得十分有用；在这种构造中它们被用于"有限域上的几何"，在此种几何中出现只有有限个点的"代数曲线"！

关于代数数域已经知道许多事情，这种域由具有有理系数的方程的一个或几个（实或复）根生成（§3，CⅢ）. 对它们的研究始于高斯，它一直延续至今成为深刻研究的主题；它是当代数

[1] 关于有限单群的完全分类，即确定有限单群的所有同构类，经过全世界上百位数学家的持久努力，于 1961—1981 年完成. 这是数学史上的一个重大成就. 最初其论证散布在 300 多篇论文之中，其中引用了许多新的群论概念，证明了大量定理. 现在这一分类已被数学家们普遍接受. 美国数学协会从 1994 年开始，出版由 D·戈伦斯坦（Gorenstein）、R·利翁（Lyons）和 R·所罗门（Solomon）撰写的拟出 12 卷的专著《有限单群的分类》（The classification of the finite simple groups, Math. Surveys & Mono., Vol. 40, American Mathematical Society）. 从已出的两卷（1994，1996）看，全部篇幅可能会有 2000 页左右. ——译注

论的主要分支之一.然而为达到探索出它所有的奥秘,仍然有长长的路径.

易于理解,当在一个结构上添加补充公理,这样得到的"比较丰富的"结构的数目会较少.拓扑域(§3,F)的情形就是如此,局部紧[①]的拓扑域都已知道.还有有序域(§3,F),它显示了很大的多样性.但如果在其公理系统上添加阿基米德公理(见第三章,§6),那么就能得到这些有序域的完全描述:每个这样的有序域是域 **R** 的一个子域;而如果我们还要求柯西区间套公设 [161] (见第三章,§6)得到满足,那么遵从所有这些公理(列举于第三章附录2中)的域仅有域 **R** 自身(在同构范围内).于是就能定义域 **R** 为仅有的有序阿基米德完全域(最后的定语"完全的"是指柯西准则满足).

C. 函子和结构的发明

面临没有能力"在同构范围内"对同一类型(即满足同样的公理系统)的所有结构进行分类,数学家们常常后退到比较粗略的分类,即把不同构的一些结构收集在一个类中,每个类中的结构常在弱于同构的意义下称为"等价的".获得这种分类的最古老的方法之一是使一个结构伴以一个数值不变量,就是说对于两个同构的结构为相等但不足以在同构范围内充分刻画一种结构的数.我们将在第六章§3,B中随同基数和维数概念给出这方面的例子;在这两个概念得到澄清前,它们曾招致某些难题.

上述想法的一种现代样式是使一个结构对应到另一类型可能不同的结构.这种对应必须遵从与函子概念相联结的条件,对后一概念此处不能详述.这些条件使得两个同构的第一种类型

---

① 一个集合 A 是紧的,如果从 A 的元素的每个序列(x_n)能选出子序列(x_{n_k}),后者在 A 中有极限.一个空间是局部紧的,如果它的每个点有一紧邻域.

结构“函子地”对应到两个同构的第二种类型结构，但反过来不一定如此. 例如，一个李群可对应出一个非结合“李代数”；或者一个可度量化空间可对应出各种群，这些群中最重要的是同调群和同伦群.

因此可以说，发明函子是当今积极追求的目标之一，而且许多例子表明，即使并非所有这些发明都产生同等重要的后果，至少我们可以期待某些发明会带来巨大的进步①.

研究具有相同类型结构的对象的一个集合或这样的对象的同构类的一个集合的另一更加精致的方法是在这个集合上考虑[162]一种新类型的结构. 最早的例子是凯莱于 1860 年发明的. 凯莱在 3 维空间中考虑所有其度②为给定的数 d 的曲线. 他证明，每个这样的曲线 Γ 对应到一个维数为 N(此数较大)的向量(§1, B)$c(\Gamma)$，它的 N 个分量遵从一定数目的代数关系，使得满足这些关系的每个向量对应唯一的一条度为 d 的空间曲线 Γ. 例如，对于直线情形($d=1$)，可取 $N=5$，而且在 5 个分量之间只有一个 2 次关系式. 对于 $d=3$，凯莱的理论表明可取 $N=399$. 我们说这就对度为 d 的所有曲线的集合 $C(d)$ 提供了一个代数簇结构. 在直线情形，克莱因应用这个结构得到“线几何”(19 世纪中很流行的一个主题)的一些新结果. 凯莱方法的推广在当代代数几何中有许多应用.

黎曼在其代数平面曲线理论中发现了某些深刻得多的性质. 他运用域论使每条不可约代数曲线对应到一个“代数函数域”，他用两条不可约曲线的代数函数域同构这个事实来定义所

① 例如通过证明一个函子使两个同一类型的结构对应到两个不同构的结构能证明前两个结构不同构. 这是证明球面与环面不同胚的方法，因为两者对应的同调群是不同构的.

② 一个曲线的度是该曲线同任意平面(在该曲线为平面曲线的情形，必须不包括它所在的平面)交点的最大数目.

给两条曲线的“双有理等价性”(附录 3,D). 更进一步,他使每条代数曲线上的代数函数域对应到一个正整数或 0,称为该曲线的亏格①g,使得两条双有理等价的曲线具有相同的亏格. 然而,反过来,两个具有相同亏格的代数函数域却不一定同构. 黎曼揭示,亏格 $g>2$ 的域同构类依赖于 $3g-3$ 个复参数,他把这些参数称为“参模”②. 到了现代,在这些类的集合上定义一个一意确定的代数簇结构成为可能,这方面的研究一直很活跃,是代数曲线论的主导课题. [163]

从那时以来,上述想法“占领”了许多别的领域: 遵从某些条件的映射的集合以及拓扑或微分结构的同构类的集合被可度量化空间、群或环等结构“组织起来”. 集合上的结构的发明初看起来没有什么,却是数学进步的另一重要源泉.

与上述想法相联系的一个概念是结构分类中的“脱殊性”. 可能发生这样的情形: 某种类型的结构往往会表现出很多捉摸不定的“病态面貌”,以至使试图分类的人望而却步. 但有时这种情况也能加以补救,办法是从给定类型的结构中选出一些可以认为是“好的”结构,它们能以合理的方式进行分类. 然而为使这种方法获得重大成果,还必须表明剔除的“坏的”结构在所研究的类型的所有结构组成的集合中是某种意义下的“小”子集. 当然,必须使这种想法精确化,它可通过几种途径达到;例如,在无穷集中,可以认为有限集是“小”的,而在赋予“测度”的集合中,可以认为测度为零的集合是“小”的,等等.

---

① 如果度为 n 的平面曲线除具有不同切线的二重点外没有别的奇点,则其亏格为 $(n-1)(n-2)/2-d$,其中 d 是二重点的个数. 方程为

$$ax^n+by^n-c=0 \quad (abc\neq 0)$$

的曲线没有奇点,因而其亏格为 $(n-1)(n-2)/2$.

② 我们已碰到过模这个名词,但意义不同;对十分不同的对象取同一名称是数学家们的一个令人遗憾的习惯.

5. 当代数学

A. 数 学 概 观

对每年发表的无数数学论著进行分类整理是一项只有超人才能完成的任务，由于它们构成紧密结合、错综复杂的网络，所以几乎不可能解开使各种理论联结在一起并成为这些理论前进主要动力的链扣．因此，下面列举的项目及简述的内容，只能看作对数学中可以加以研究的各种方向以及关于这些方向中看来常常未能显示其倾向的问题的非常大略的提示．

Ⅰ）形式逻辑和集合论

虽然莱布尼茨（在其很长时间一直未发表的著作中）可以看作这方面的先驱者，但是在大约 19 世纪中叶以前，数学家一直没有探讨这些科目．在第六章中，当讨论数学"基础"问题，它们的起源，它们在 20 世纪中的可观发展以及它们相对数学其他领［164］域所处的地位时，我们会回到这个论题．必须注意，最近发现，在计算机科学中，在各种类型的"自动机"理论中，以及在可以冠以"信息论"的所有问题中，形式逻辑具有完全出人意料的应用领域．

Ⅱ）组合论

首先，可以说这个名称指的是集合论的一个部分，即有限集理论以及计算按各种方法构造的有限集的元素数目的问题．某些这类问题例如计算 n 个元素的集合中两个元素构成的偶的数目（它等于 $n(n-1)/2$），可追溯到古典时代．我们遇到的另一例子是计算 n 个元素的集合的排列数 $n!$（§1，D）．数学家们在很长时间中多少有点轻视属于这类研究的方法，他们喜欢把它归入"数学游戏"，即使得通俗科学刊物的广大读者感到兴

趣的那类数学.但在最近几十年中,情况大有改观;人们注意到"组合论"方法在完全"受人敬重"的数学领域诸如群论、李代数论、代数几何以及代数拓扑中可以有重要应用."组合论"方法也成功地应用于数学的各式各样应用(包括"信息论")中.

Ⅲ) 范畴和函子

这是数学中最新领域之一,诞生于1943年,来自艾伦伯格和麦克莱恩关于代数拓扑的探索.这里不可能谈论"二阶抽象"能考虑些什么;在这种抽象中,赋予结构和相互之间的映射的集合消失,更正确地说"升华"为与"总合"成"法则"这些日常概念毫无关系的"对象"和"箭矢".这一理论先是在代数拓扑中十分有用,然后在代数几何中成为具有根本重要性的工具.它推进到数学中别的许多领域,最终成为自主发展的对象.

Ⅳ) 代数学

关于基础代数结构的一般概念——群、环、域和模,现在已是大学一年级水平代数教学的一部分.当前代数学研究的较大部分集中在更专门的代数结构上.这些结构最初被看作简单的例子,它们不久前才得到开发,但是这个领域中的方法和结果繁殖到这样一种程度:其中每个结构的研究分开来看都已成为一 [165] 门主要的理论.这就是下面Ⅴ)到Ⅷ)所列举的结构.在"代数学"这个总标题下所剩的是称为"泛代数"的科目,其合成律并不遵从诸如定义群或环结构那样具有限制性的公理.

至于"经典"代数即本质上是代数方程组的研究,如今是交换代数(第Ⅵ项)的一部分.

Ⅴ) "抽象"群

这一理论始于若尔当所著"Traité des substitutions et des equations algebrique"(1870)(《置换和代数方程专论》),它已具

有自己独特的方法和问题,并已成为一个庞大的分支.同样毋庸置疑的是它在数学所有领域中具有最多的应用,其程度有如人们所说:当你不很了解新的数学对象的性质,就应当试着在其上设置一个群结构.看起来这好像是心血来潮,但事实上却已取得不止一次的成功.最接近于"抽象"群的数学领域是李群、代数几何、代数拓扑和微分拓扑以及数论;以这些结构为中介,群论使得数学中其他许多问题及其物理应用丰饶多产.

Ⅵ)结合非交换代数和非结合代数

这是可以追溯到19世纪中期的某些理论.它们涉及结合的、非交换的环(诸如四元数域(§2,D))或非结合环(其中乘法不是结合的,但保留关于加法的左、右分配性).人们力图对它们进行分类,通常通过设置补充公理.它们已进入"抽象"群论、李群论、代数拓扑和泛函分析中.

Ⅶ)交换代数

这是关于交换环和交换域的研究,始于约1860年,也已成为代数学的一个独立分支.它也有自己的问题和方法,最重要的是,现在它同代数几何有紧密的共同生长关系.它向代数几何提供专门工具即交换环的性质,但是近来它转过来得益于自然地属于代数几何问题的直观考虑,因为现在有可能把交换代数问题翻译为格罗腾迪克概形论中的几何术语,它们是代数曲线和[166]代数曲面概念的终极推广.

Ⅷ)同调代数

它涉及这样的函子的构造,这些函子使每类范畴的对象对应到另一些对象,最常见的是交换群.这种理论与代数拓扑相联系,与范畴和函子理论(第Ⅲ项)同时,大约诞生于1943年.它的应用领域不断扩展到几乎整个数学.可惜由于它所用概念极端

抽象，此地不可能对它进行更多的说明.

Ⅸ）一般拓扑学

这个数学领域的肇始可定于给出度量空间定义（§3，E）的1906年．如同代数学那样，最常用的拓扑结构（可度量化空间及其最简单的推广）的公理和一般性质如今已是大学一年级课程的教学内容．在这里，代数拓扑和微分拓扑（第Ⅹ项）以及函数空间（第Ⅻ项）的主要理论也迅速地从这些一般考察中分离出去．正在进行的研究涉及其他数学领域几乎不用的拓扑结构.

Ⅹ）代数拓扑和微分拓扑

在同构范围内对所有拓扑结构加以分类是一项不现实的任务．这方面的努力很快限于最接近我们的几何直观的空间，诸如曲面及其推广——“有限维簇”，后者在泛函分析（第Ⅻ至ⅩⅥ项）及其物理和力学应用①中是不可或缺的．通过各种几何性质的构造来扩展这一领域的途径已经找到．如前所述（第Ⅷ项），对所有这些空间都能定义使一个空间对应到一个代数结构的函子，而代数拓扑和微分拓扑就是这些函子的研究．这个分支创立于1900年，在其后50年中它的发展是爆炸性的．或许这是数学中 [167] 出现崭新概念最多的领域，其中许多概念在乍看起来距离极远的理论中得到出人意料的反响．它构成一个不断更新的壮丽大厦，其复杂程度使得极少有专家能全部掌握.

① 例如，空间中固体的位置依赖于6个参数，因此可由6维空间中的一个“点”标志．为看出这一点，例如考虑空间中立方体的位置．首先是它的中心必须固定，中心可由它的3个坐标给出；然后，经过中心垂直于立方体某个面的垂线（“四分旋转轴”，见附录2，E）依赖于2个参数，它可由该轴与中心为立方体中心、半径为1的球面的交点的坐标来实现．最后，当这个轴固定后，仍能使所给立方体绕该轴转动，此转动再给出一个参数，即转动的角度.

XI）经典分析

我们曾提及(§1,A)——但未加精确描述——19世纪数学令人击节赞赏的最美花朵之一，即由柯西、黎曼和魏尔斯特拉斯开发的关于单复变量解析函数的伟大理论. 时至今日它继续吸引着大批数学家，它也得益于从代数拓扑和泛函分析移植过来的技巧.

在19世纪后30年，随着康托尔的工作，还开发了分析的一个分支. 在这个分支中，有意把解析函数置于一旁，而把目标定在探索对于比解析函数多得多的"不规则"函数(见第六章，§2,B)，还能说出些什么正面性质. 这里涉及的是任何个数实自变量的函数，更恰当地说，如今这方面的研究依赖于拓扑学.

经典分析长出许多新方向，其数目比数学中任何别的领域为多，下面第XII至XIX项阐述其中最重要的分支.

XII）积分和概率论

我们看到(第三章，§9)，17世纪出现了实变量的实值连续函数在一个区间上的积分的定义. 在18世纪，类似地出现了定义于$\mathbf{R}^2$的子集D上的两个实变量的实值函数$f(x,y)$的二重积分的定义. 如果$f(x,y)\geqslant 0$，则积分

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$$

是$\mathbf{R}^3$中由关系

$$(x,y)\in D,\quad 0\leqslant z\leqslant f(x,y)$$

确定的体积.(例如，如果取D为圆盘

$$x^2+y^2=R^2,$$

取$f(x,y)=\sqrt{R^2-x^2-y^2}$，则所说关系确定一个半球体.)在18世纪也有可能定义任何个数的实变量的实值函数的"多重积分"，尽管此时这种积分不再有任何几何直观形象.

从 1895 年至 1930 年,人们认识到积分概念还能进一步远远扩展.现在它可施之于定义在一个集合 E 上的非常一般的实值函数,E 也可以非常任意,只要假定能对 E 的某些子集 A 定义一个测度,即对每个这样的子集赋予一个数 $\mu(A)\geqslant 0$,它是长度、面积和体积这些经典概念的推广,并满足少数几条公理. [168] 公理之一是对两个没有公共元素的集合 A,B 之并 $A\cup B$,有

$$\mu(A\cup B)=\mu(A)+\mu(B),$$

即此测度是加性的.

能在比经典分析广泛得多的意义上定义测度和积分带来了巨大的进步,尤其是在泛函分析和李群理论中.加之,它赋予概率论以新的生命.这门学科在 17 世纪由帕斯卡和费马创立,几乎一直完全限于机会博弈的理论;尽管雅格布·伯努利和拉普拉斯使它取得了显著进步,它仍然缺乏坚实的基础,而把它扩展为所谓"几何概率"的尝试并未取得多大成就[①]. 1930 年,柯尔莫哥洛夫揭示,通过把概率论同测度论联系起来,就有可能为概率论奠定基础:所必须的是只考虑取值于 0 与 1 之间的测度[②].沿着这条途径决定性地并入数学后,概率论成为全速扩展的一个领域,它对其他科学的无数应用已为人熟知,不必再在这里予以强调.

Ⅷ) 函数空间和算子

从 17 世纪起,几何学、力学和天文学中的问题(例如牛顿从万有引力定律推导开普勒定律)导致"函数方程",即方程中的"未知量"不再是数,而是函数.稍后(第 XV 和 XVI 项)我们要谈到这种方程的最早类型,即常微分方程和偏微分方程.

---

① 不同的作者按照他们各自关于这种概率的概念,可能对同一事件给出不同的概率值.

② 整个集合的测度必须等于 1,而空集的测度必须等于 0.

在 20 世纪初期,随着度量空间观念(§3,E)的出现,下面的想法变得平常:正如具有 n 个未知数的代数方程组的解可以看作 $\mathbf{R}^n$ 中的点(或向量),作为函数方程的解的函数也可看作某个度量空间中的点.为能使用这类语言,我们必须说明两个函数之间的距离(§3,E)是什么;很快就清楚,定义这样的概念有许多不同的途径(附录 4).

一当这样定义"函数空间",下一步就是考虑函数空间 E 到[169]一个空间 F 的某个映射(§3,B),它常称为算子.类似于实变量的实值函数,算子能以各种方式组织起来(§1,C),因而能构成群和环;还有可能定义关于算子的积分,等等.经典物理学已能应用这些"抽象"理论(第XIV项).随着量子力学的发展,这些理论变得十分重要,在量子力学中,物理量(位置、动量、能量等等)解释为希尔伯特空间上的算子.

我们也必须提到正在迅猛发展的最优控制理论,它是变分法的现代形式,它的所有研究都在函数空间中进行.

XIV) 交换调和分析

这一理论的来源是毕达哥拉斯学派关于张紧的弦振动发出的声音可"分解"为一个"基音"的"泛音"这个发现,其数学理论起源于 18 世纪.它的主要因素是把周期函数 f(即对每个 $x\in\mathbf{R}$ 满足 $f(x+2\pi)=f(x)$ 的函数)"分解"为"简单"周期函数

$$a_n\cos(nx+\theta_n)$$

之和.这种分解一般含有无穷多项,即它是一个级数.这方面的深入研究由傅里叶于 1807 年开始,嗣后从未间断地吸引着分析学者,并在物理学中有无数应用(附录 5).

紧接"傅里叶级数",傅里叶通过创造所谓"傅里叶变换"扩展了他的研究范围.这种变换是算子(第XIII项)的最早例子,它建立了定义于 $\mathbf{R}$ 上的实值函数(一般不是周期的)与另一也是定义于 $\mathbf{R}$ 上的函数之间的对应.这种算子及其推广在偏微分方程

(第XVI项)、数学物理和概率论中是非常有用的工具.最近注意到,在这种理论同可度量化交换群(§3,F)理论之间有着紧密联系,而调和分析已占领了代数数论(第XXI项).

XV) 微分方程

从18世纪以来,除极少数情形外,并不知道如何显式地写出微分方程的解[①].除了由弦或杆的振动问题引起的方程外,在 [170] 很长时间中微分方程组的一般理论限于依据所给方程在一个点的邻域内的性态,对解进行*局部*研究.是庞加莱于1880年开始首先成功地得到涉及解(它常称为轨道,因为自变量 t 常被同化为"时间")的整体的定理.如今这种理论界定为*动力系统*理论,当然它在力学和天文学中有众多应用.它在19世纪得到巨大扩展,在其发展进程中越来越同代数拓扑(也是由庞加莱使用的)和测度论(在来自物理学的称为"遍历的"问题中)紧密相联.

XVI) 偏微分方程

最早一批偏微分方程与微分几何学(第XVII项)和物理学问题相联系,出现于18世纪.但由于缺乏合适的数学工具,19世纪前不可能对其一般理论开始进行研究.开始时设想,当从一个自变量转向几个自变量时,用于研究常微分方程的方法易于推广到研究偏微分方程.但很快认识到,两种理论之间的相似性是非常表面的,由偏微分方程提出的问题需要完全新的探索方法,其研究有时导致惊人的结果,例如1956年发现存在没有解的偏微分方程.自1950年以来,从泛函分析和傅里叶变换(第XIV项)理论发掘的新概念使偏微分方程理论有可能出现令人惊异的新转折,它现在正全速扩展,并且仍然与物理应用紧密相联.

---

① 刘维尔能证明,诸如 $xy' = e^x$ 这样只出现"初等"函数的微分方程,其解却不是"初等"函数.

XVII) 微分几何学

从 17 世纪起,数学家们把微分学应用于研究平面曲线的局部性质,诸如确定一点处的切线,确定拐点和多重点,定义曲率.这些结果在 18 世纪推广到空间曲线和曲面.然后黎曼于 1854 年大胆地使任何维数的"流形"概念化,并对它们定义作为早先概念推广的微分概念,尽管它们已不再像曲线或曲面那样能由"具体形象"给出.黎曼的继承者可观地发展了他的想法,并以同代数拓[171]扑较前更加密切的联系着手研究整体问题.这一看来远离自然科学的理论,在广义相对论和宇宙论(在这些理论中 4 维或更高维空间的观念起着基本作用)中具有完全意想不到的应用.

XVIII) 解析几何学

19 世纪末期,当数学家们希望把柯西、黎曼和魏尔斯特拉斯得到的关于单复变量解析函数的深刻结果推广到多复变量解析函数时,他们遭遇到始料未及的困难.直到 20 世纪中期,部分由于使用来自代数拓扑的新概念("层上同调"),这些困难才得以克服.从那时以来,多复变量解析函数论也出现了更加几何化的转变,更加接近代数几何.必须紧挨"代数簇"研究"解析簇";例如,两个复变量的方程

$$F(x,y)=0$$

(其中 F 不再是一个多项式而是一个解析函数)定义一条"解析曲线".因此如今把多复变量函数论称为解析几何学[①].

XIX) 李群

由 S·李以"连续群"(§3,F)名称引进的可度量化群在很长

① 从 18 世纪末以后,"解析几何"一词用来指坐标方法用于欧几里得几何.如今除初等教育外已不再这样使用这一名词.

时间内在其他所有数学理论中占有边际位置. 由于 E·嘉当和 H·外尔于 1910 至 1930 年间所做的工作,数学家们开始意识到它们在微分几何中的重要性. 以后它们可以应用的范围不断扩大,如今它已遍及整个数学,从"抽象"群论和数论一直到"基本粒子"物理的数学模型. 由 H·庞加莱开创的自守函数论在这个科目中占有中心地位.

XX) 代数几何学

虽然古希腊人已研究过某些代数曲线和代数曲面,然而直到笛卡儿和费马发明坐标方法后,这些对象的一般研究才得以展开. 19 世纪以来,复值坐标和"射影"概念的引进给予较少依赖于使用坐标的想法以新的生命. 这些想法取得了引人注目的 [172] 成就,使其理论在数学家中十分流行. 同时黎曼通过他关于代数函数的深刻结果,使代数几何学回到同复变量解析函数论和同拓扑学的联系上. 这些联结比过去产生了更多成果,但从大约 1930 年起,能几乎全部把使用分析的证明代之以仅依赖于交换代数的另外的证明. 它为这一理论带来巨大扩展,并使它接近数论,并为在这两种理论中同样有用的问题提供想法.

XXI) 数论

无疑这是最古老的数学理论,而对很多数学家它仍然是——引用高斯的话——"数学的女皇". 数论中的问题,由于它们在陈述时明显而又简单,而为解决它们又必须克服极大的困难(参见第四章),因而总使研究者着迷. 它总是成功地求助代数和分析,如今又求助拓扑学. 它被认为是"最纯粹的"数学学科,然而最近开始有某些应用.

B. 专才和通才

一个在大学学习了三四年获得了必不可少基础知识的年青

数学工作者,如果希望成为上面列举的 21 个领域中某个领域的专家,仍然必须研读许多书籍和大量论文,才能使自己与所选领域中的技巧和问题并肩前进.因此大多数数学家终其一生限于这些理论中的某一项,甚至某一理论众多方面的一个方面,就毫不足怪.某些数学家通过拓广已知的技巧使之超越前人所能做到的程度,设法使这些技巧获得足够的威力以超越他们的前辈未能越过的障碍;这些人可比作战术家,他们以凶狠的搏斗攻克坚固设防的堡垒.但是也会有几位富有才智的人物,他们能精通看起来相距遥远的几个领域,而且能洞察把一个领域的结果用于解决另一领域中问题的道路;他们就像战略家,凭着聪明的计谋几乎毫不费劲地把号称不可攻破的阵地夷为平地.最近的一个例证是这方面的典型:通过运用某些直接滋生于基本粒子物[173]理学的偏微分方程的性质,年青的英国数学家 S·唐纳逊找到了关于 4 维微分流形拓扑问题的答案;这些问题曾经难倒这方面的专家,他们从未想象到有可能以这种方法处理它们.

在最近 40 年中,几乎每年总有新人来临,他们那样成功地攻克了几十年中未能解决的问题.没有任何理由认为不断涌现天才的潮流注定会干涸,任何时代数学都不像我们这个时代这样繁荣昌盛.

C. 数学理论的演变

熟知在自然科学中,观察和实验组织为阐释体系("理论"),它从少数多少易于观察的实体以及认定能描述这些实体的行为的定律(例如关于"质点"力学的牛顿定律)出发,逻辑地推导出所观察的现象将如何展开.理论与观察之间的对应总会被误差玷污,误差可能来自测量仪器不够精确.只要它们不太严重,就可相安无事.但是,当误差如此巨大以致理论不再能解释实体行为时,已有的理论就必须加以修正,如果有必要,就得完全推翻它的基础.如所周知,在过去 80 年中,这样做过若干次.

数学的情形完全不是这样．一旦某个公理系统中的一条定理其证明被公认为正确，它就不会再带来疑问：欧几里得的定理在今天确实有效，就如同2300年前一样．这是否意味着数学没有演变？本章和上章所说的一切表明情况正好相反．然而这种演变并不仅仅在于新定理的积累；因为新定理并不是简单地添加在旧定理上，而是吸收旧定理，把它们转变为“推论”，有时最后甚至除了历史学家会提到外，再也不会有人明显地谈到它们．甚至定理的表述也会完全改变．例如，作为古希腊数学巅峰的关于正多边形的描述，如今表达为3维空间位移群的有限子群的描述（§2,B）．其他的转变或许显得更加奇特：一个公理系统被另一个“等价的”系统[①]代替；一个陈述，在旧系统中是一条定理，在新系统中却变成一条公理，而旧的公理则变成定理[②]． [174]

因此，两代之间发生的数学概念改变是一种重新组织，它关注新增添的内容及其与旧定理的关系．自从欧几里得起，这种变化总是在系统阐释的著作中得到巩固．从19世纪初叶开始，这种著作主要是供不同水平大学生使用的教科书．科学史专家通过比较这种著作，可以了解某个时代数学界成员的思考方法，而不只是主导那个时代的天才的想法．天才的思想几乎总是超越他的同时代人，预言下一代会普遍接受的观念．现在数学理论是按刻画它们的结构分为一些如上面所列举的粗略的范畴来详尽阐述的，不过这样做只不过约有50年，而从历史上看，正如我们试图表明的，关于这些结构的概念及其运用要追溯到19世纪和20世纪前三分之一年代．

如果说从16世纪以来数学以前所未有的速度持续不断地前进，那么首先应归功于数学家从不间断地提出各种各样的问

---

① 两个公理系统是等价的，如果从其中之一推导出来的全部定理也能从另一个推导出来．

② 某些性格保守的数学家从未停止反对这种做法．

题.希尔伯特说:"只要一门科学提供大量问题,它就是鲜活的.缺少问题意味着它的死亡或停止发展."

对于A中列举的大部分分支,从它们存在之日起,与其发展须臾不可或离的问题之流从未干涸过,不管它们是数学应用提出的问题,还是像我们在第四章中举过例子的来自数学"内部"进化的问题.

但是也有一些例外.我们在本章开头所指出的分析进展中于1785至1795年间出现的奇怪的空档,或许是由于感到没有能力处理当时出现的物理问题.解决这些问题需要发明新的方法.

另一个在光辉灿烂的繁荣后衰落随之而来的例子是位于代数与代数几何交界处的不变量理论.在半个世纪(1840—1890)[175]的大量论著以它为主题之后,它所提出的最重要问题为希尔伯特全部解决.由于缺乏新问题,这一理论休眠了50年.只是到了最近,由于它同李群论(第XIX项)和代数几何中黎曼"参模"理论(§4,C)的联系,这方面的研究积极性才又极大地迸发出来.

我们多次强调,数学的进步可能源自表面上十分不同的理论汇合于一个问题;然而也能出现分散的现象.例如,我们谈到过一般拓扑学(第IX项)诞生的动力是为更有效地处理泛函分析中的问题提供结构.有可能如同哈密顿所做的那样(§2,D),改变这些结构的公理,研究由此得到的新结构的性质.但除了极少数情形外,它们在研究算子或函数方程中没有进一步的应用,以致最后结果是它们成为数学中两个不再相互接触的领域.同样情形出现于代数学中,对于它作为"消费者"的理论主要是代数几何学、李群论和数论,而属于第VI和VII两项的许多研究倾向于分裂,达到几乎完全分离的程度.并不是每个人都享受得到哈密顿(身后)的幸运.

6. 直觉与结构

所有伟大的数学家在谈到他们的工作时都乐于坚持在其研

究中起作用的是他们通常称为的"直觉"①.未被传授现代数学知识的人可能会深感奇怪：如果他打开今日的一本数学书,那么除了成百上千的引理、公式、定理和推论外,他会什么别的都看不到;而这些东西又按残酷无情的逻辑规则以复杂的方式联结起来,与之有关的是在我们的感觉世界中没有任何"形象"的数学对象.我认识一些老数学家,他们无可争辩地是经典分析的大师,但不能想象他们的年青一辈如何能在"抽象"的大海中应付裕如.这些年青人乐意把他们的推理方式比作机器运转,他们操纵公式而不试图理解这些公式. [176]

我相信真理不能再前进半步;但是显然我们一定不能在通常意义下理解"直觉"这个词.困难在于,数学家所说的"直觉",是一种难以言传的完全个性化的心理体验;而且有充分理由认为两位数学家的"直觉"往往大相径庭.

然而,指出这些心理体验的某些我认为是相当普遍的特征,尽管它们主要基于我个人的回忆,或许并非全是虚幻的目标.首先,当你开始对在此之前你一无所知的一种理论感到兴趣时,你对它毫无直觉,尽管你能一步一步地验证属于这一理论的全部定理.你提出稍后看会是遥不可及的问题,你完全不能用你自己的类似于你读到的论证来思考.然后,如果你坚持下去,面纱就会一点一点地揭开,你开始理解对这一理论作出贡献的数学家为什么沿这条途径而不沿另一途径前进.你对该理论所讨论的对象熟悉起来,你认识到它们有"自然的"行为方式,必须仔细不要去弄乱②.正是在这个时刻,你或许会幸运地邂逅以前所没有

① 盎格鲁一撒克逊人更喜欢用"洞察力"(insight)这个词,它或许比具有神秘气氛的词更合适.

② 例如,当两个结构为同构(§4,A)时,可能存在从其中一个结构到另一个的某些同构,它们比别的同构更"自然".使得这个多少有点模糊的想法精确化的尝试导致"函子"概念.

一具有更神奇性质的"蒂茨厦",它是富有成果的应用的源泉,而没有几何语言很可能根本想不出它.

因此很容易理解,数学家如此欣赏这种语言的优点,以致要不了多久他们就能很好地超越这些最初推广的界限.经典线性代数处理数组(§1,B)和矩阵(§2,B),而在几何语言中它们成为向量和线性映射.然而,现在这种语言已扩张到这种程度,其中代替数的"标量"是任何域甚至任何环(§3,CⅡ)的元素,它具有使得论证比较简短的功效,以致如今当我们而临例如线性[179]代数的最新推广——同调代数(§5,A,Ⅶ)时,甚至不能看出没有这种"空间"语言怎么能进行下去.

通过引进由系数取自任一域的方程定义的簇的研究,同时保留该域为实数或复数域情形沿袭下来的语言,经典代数几何以同样方式得到推广.每逢能不用分析构造证明时,这样做总是可能的.

几何语言也在另外方向上推广.我们解释过(§5,A,Ⅻ)函数如何变得可以说成函数空间中的"点".这种几何语言导致出人意料的富有成果的联系.例如,位势理论中一个困难的经典问题——"狄里克雷问题",解释成到直线或平而上的正交投影这个经典几何算子在适当的"希尔伯特空间"中的推广.但这还不是这种通向"拓扑"概念的新一阵进攻的终点,它们甚至还侵入数论,在那里我们如果能使自己处于一个点的邻域中,就能谈论"局部"性质.这样我们看到,现在可以毫不犹豫地肯定,几何精神比以往任何时候更强烈地支配着整个数学躯体.

我们强调过这种观念"转移"来自基础教育中得到训练的几何推理,因为它无疑是最引人入胜和最广泛的.然而,也有许多来自代数、分析或数论的类似的转移.这是当代数学的主导特[180]征——统一性(对此不能强调过分)的标志.

尽管它已不再涉及可感知的现实性. 这样,我们谈论"n 维空间中的点"来代替 n 个数的数组(§ 1,B),谈论"超平面"来指明满足线性方程

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n + b = 0$$

的"点"$(u_1, u_2, \cdots, u_n)$的集合,其中系数 $a_j(1 \leqslant j \leqslant n)$是实的且不全为零;等等. 对于力学中提出的问题,变量一般还有一些关系联系起来,于是我们必须只考虑位于满足这些关系的子集上的点$(u_1, u_2, \cdots, u_n)$,这样的子集通常称为"簇",如果有 p 个关系,则一般称该簇是"$n-p$ 维"的①. [178]

这样,从 19 世纪后期起,早先局限于通常 3 维空间的代数几何和微分几何的研究扩充到容纳任何维数的簇和流形;而从庞加莱时代开始,就认为簇和流形构成数学分析必须发展的领域(它也称为"整体分析"). 经验证明,3 维几何中的许多结果在这个扩充了的框架中有"自然的"推广. 举一个简单但多少有点生硬的例子:通常空间中一条直线与一张曲面"一般"相交于一个点有如下推广: 在 n 维空间中一个p 维簇与一个 $n-p$ 维簇想必有"一般"由孤立点组成的交. 这种由语言相似性指引的从通常的"几何直观"到新的"抽象直观"的转移已被证明具有相当大的功效,不去使用会是很可惜的. 用几何语言代替代数语言几乎总能做到相当的简化,并使掩埋在一大堆错综复杂计算中未被觉察的性质显现出来.

当线性规划和最优化数学模型中出现的不等式用多维空间中凸集性质解释时,这种模型就易于理解并能处理得好得多. 半单李群的分类最初是用大量行列式计算来做的,现在则基于几何格局——"外尔房"的研究,它的神奇性质使经典几何学中最优美的结果都显得暗淡失色. 从这种格局出发,最近还发明了另

① 当然这种语言并不改变定理的实质,它们总得通过代数演算来证明. 新的语言至多提供较简短和较有表现力的陈述.

的一条新定理或一种新方法."直觉"来到,但光靠直觉不够:你所窥见的证明必须遵守铁面无私的逻辑规则,它们必须主宰证明的各个部分.这是痛苦的劳动,肯定不时遇到未曾料想和令人沮丧的挫折.

当他们处理的概念完全没有感官可以感知的"形象"时,当代数学家怎样能够从事这种新发现的探索?无疑情形在各个时代都相同,就像我们在例如哈密顿那里(§2,D)看到的那样.我相信这些数学家为自己创设了这些数学对象的纯粹心智的、不可言传的形象.精确表述定义这些对象的公理可能有助于形成这些形象,这是由于这种表述消除了可能在这些对象的结构的各种应用中附加到对象上的所有不必要的特殊性质;尽管看起来这是自相矛盾,其实抽象有助于形成"直觉"而不是使它瘫[177]痪.

我觉得,甚至更加富有成果的是可以称之为直觉转移的过程.正如我们几次提到过的,单个结构(例如向量空间结构(§3,F))在一些定理中都被发现.每一次该结构都同另一结构中不能出现的理论的特殊性质相联结.必要时适当修改后一结构的语言,可能发生也把前者的特殊性质带给后者的情形,有时这就能产生新的非常丰富的"直觉".

这类转移的典型例子是几何语言传播到整个数学.在笛卡儿和费马发明把几何问题归结为代数问题的坐标方法(见第三章,§8)之后不久,就注意到反过来,涉及变量不多于2或3个的代数问题,如果把变量看作点的坐标,那么它也能解释为与之等价的几何问题.这种翻译能强烈地抓住从几何经验中汲取的直觉,可以用来指引问题的解决的直觉.

可惜,当变量数目超过3时,这一几何解释就不再出现.而从18世纪起,尤其在力学和天文学中,这个数目(所研究的系统的"自由度"数)可以任意地大.从19世纪中叶以来,数学家们决定使用从经典几何那里复制过来的习惯语言来处理这些问题,

附　　录

1. 四次方程的解

如果 x_1, x_2, x_3, x_4 是方程

$$t^4 - a_1 t^3 + a_2 t^2 - a_3 t + a_4 = 0$$

的根,则我们已看到(§1,D),$s_1 = x_1x_2 + x_3x_4$, $s_2 = x_1x_3 + x_2x_4$ 和 $s_3 = x_1x_4 + x_2x_3$ 是 3 次方程

$$t^3 - b_1 t^2 + b_2 t - b_3 = 0$$

的根,其中 $b_1 = a_2$, $b_2 = a_1a_3 - 4a_4$, b_3 是 a_1, a_2, a_3, a_4 的一个多项式,我们不予写出.为计算 x_1, x_2,我们来计算 $x_1 + x_2$ 和 x_1x_2.

首先,$x_1x_2 + x_3x_4 = s_1$, $x_1x_2x_3x_4 = a_4$,因此 x_1x_2 是方程

$$t^2 - s_1 t + a_4 = 0 \tag{1}$$

的一个根.

其次,让我们考虑多项式

$$y_1 = x_1 + x_2 - x_3 - x_4.$$

由于 $x_1^2 + x_2^2 + x_3^2 + x_4^2 = a_1^2 - 2a_2$,我们有

$$y_1^2 = a_1^2 - 2a_2 + 2(s_1 - s_2 - s_3) = a_1^2 - 4a_2 + 4s_1,$$

于是 y_1 可通过取平方根得到.

最后,由于 $x_1 + x_2 + x_3 + x_4 = a_1$,我们有

$$x_1 + x_2 = \frac{1}{2}(a_1 + y_1).$$

如果 α 是(1)的一个根,则 x_1, x_2 是

$$t^2 - \frac{1}{2}(a_1 + y_1)t + \alpha = 0$$

的根.于是有 $x_3x_4 = s_1 - \alpha$, $x_3 + x_4 = (a_1 - y_1)/2$,从而 x_3, x_4 是

$$t^2 - \frac{1}{2}(a_1 - y_1)t + s_1 - \alpha = 0$$

的根.

2. 关于群与代数方程之解的补注

A. 对称群 $\mathfrak{S}_n$

设 E,F 都是有 n 个元素的有限集. 通过对 n 进行归纳推 [181] 理可知一一映射 $f:E\to F$(§4,A)的数目等于 $n!$. 事实上,设 a 是 E 的一个元素,b 是 F 的任一元素,则对从 $E\setminus\{a\}$ 到 $F\setminus\{b\}$ 上的每个一一映射 g,恰对应出 E 到 F 上的一个一一映射 f,使对 $x\in E\setminus\{a\}$ 有 $f(x)=g(x)$,且 $f(a)=b$. 由归纳法假设,这给出 $(n-1)!$ 个一一映射. 由于 F 中选取 b 有 n 种方法,故按阶乘定义从 E 到 F 上的一一映射个数为 $n\cdot(n-1)!=n!$. 如果取 $E=F=\{1,2,\cdots,n\}$,则从 E 到其自身的一一映射是柯西的"置换",而其数目为 $n!$.

B. 方程的伽罗瓦群

与其前人相比,伽罗瓦的独创性在于对每个没有重根、系数属于 $\mathbf{C}$ 的任一子域(§3,CⅢ)K 的方程 $P(t)=0$,对应一个由置换构成的群,通常它与对称群不同. 设 $z_1,z_2,\cdots,z_n$ 是 $P(t)=0$ 的不同的根,因之

$$P(t)=a_0(t-z_1)(t-z_2)\cdots(t-z_n).$$

$\mathbf{C}$ 的由 $z_1,z_2,\cdots,z_n$ 生成的(§3,CⅢ)子域 $E=K(z_1,z_2,\cdots,z_n)$,称为 P 的根域. 我们记得 E 的元素是系数在 K 中的有理分式

$$\frac{R(z_1,z_2,\cdots,z_n)}{S(z_1,z_2,\cdots,z_n)} \tag{1}$$

的值,其中 $S(z_1,z_2,\cdots,z_n)\neq 0$. 对于每个置换 $\sigma\in\mathfrak{S}_n$,假定 $S(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))\neq 0$,则数

$$\frac{R(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))}{S(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))} \tag{2}$$

属于 E. 但可能发生这样的情形:对于多项式 R,有

$$R(z_1,z_2,\cdots,z_n)=0 \text{ 但 } R(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))\neq 0.$$

满足下述条件的置换 $\sigma\in\mathfrak{S}_n$ 的集合 G 称为方程 $P(t)=0$（或多项式 P，或域 E）的伽罗瓦群：对每个具有性质 $R(z_1,z_2,\cdots,z_n)=0$ 的多项式 R，恒有 $R(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))=0$. 很清楚 G 确是 $\mathfrak{S}_n$ 的子群，因为如果 σ,τ 属于 G，则由 G 的定义，复合 $\sigma\tau$ 同样属于 G.

例

取 $K=\mathbf{Q}$，考虑方程

$$t^4-2=0. \tag{3}$$

由于 $i^4=(-1)^2=1$，故此方程的 4 个根是

$$z_1=\sqrt[4]{2},\ z_2=i\sqrt[4]{2},\ z_3=-\sqrt[4]{2},\ z_4=-i\sqrt[4]{2},$$

而且在这些根之间有许多有理系数多项式关系，例如

$$z_1z_2-z_3z_4=0.$$

这就易于给出 $\mathfrak{S}_4$ 中的置换但不属于相应的伽罗瓦群 G 的例子，例如循环置换 $(z_2z_3z_4)$（§2,B），因为 $z_1z_3-z_4z_2\neq 0$.

可以证明 G 由两个循环置换

$$\sigma=(z_1z_2z_3z_4),\quad \tau=(z_2z_4)$$

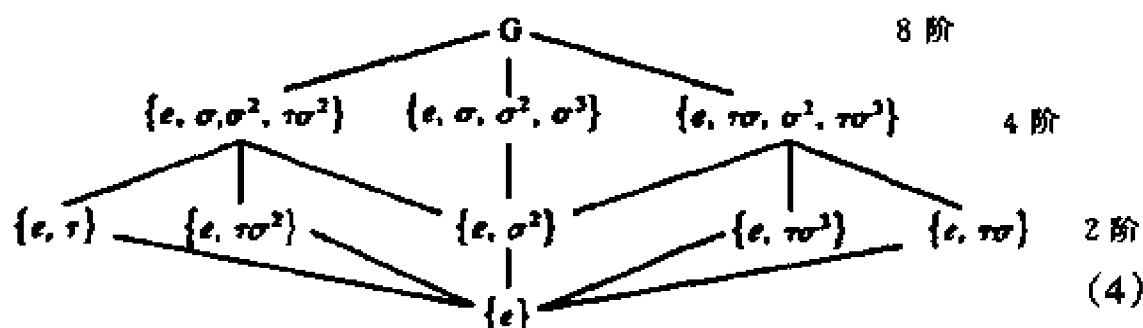
生成，σ,τ 满足 [182]

$$\sigma^4=e,\ \tau^2=e,\ \sigma\tau=\tau\sigma^3.$$

这表明群 G 不是交换的. 群 G 有 8 个元素，即

$$e,\sigma,\sigma^2,\sigma^3,\tau,\tau\sigma,\tau\sigma^2,\tau\sigma^3.$$

G 的子群系按包含关系排序，可由图表(4)容易地描绘.



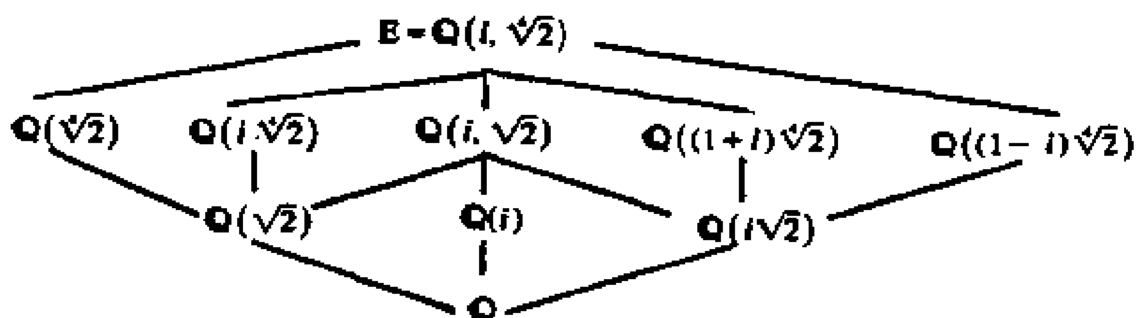
(4)

C. 伽罗瓦群和自同构群

根域 E 的伽罗瓦群 G 运算于 E 上(§3,F).如果 x 是 E 的元素(1),则 $\sigma\cdot x$ 是元素(2).当然必须验证,如果 E 的一个元素写成两种方式 $R(z_1,z_2,\cdots,z_n)$, $S(z_1,z_2,\cdots,z_n)$, R,S 是不同的有理分式,则仍有

$$R(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))=$$
$$S(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n)).$$

但这等价于(令 $H=R-S$)关系式 $H(z_1,z_2,\cdots,z_n)=0$ 蕴涵 $H(\sigma(z_1),\sigma(z_2),\cdots,\sigma(z_n))=0$,而这正是群 G 的定义.显然 $e\cdot x=x$, $(\sigma\tau)\cdot x=\sigma\cdot(\tau\cdot x)$.特别是,$\sigma^{-1}\cdot(\sigma\cdot x)=e\cdot x=x$,这表明 $x\longmapsto\sigma\cdot x$ 是从 E 到其自身上的一个一一映射.此外,我们有 $\sigma\cdot(x+y)=\sigma\cdot x+\sigma\cdot y$, $\sigma\cdot(xy)=(\sigma\cdot x)(\sigma\cdot y)$;换言之,$x\longmapsto\sigma\cdot x$ 是从域 E 到其自身上的一个同构.这样的映射称为 E 的自同构.伽罗瓦理论的基本定理是,存在从 G 的子群的集合到 E 的子域的集合上的一个一一映射,使得 $K\subset L\subset E$.对应于 G 的子群 H 的是 E 中在 H 的作用下不变的元素组成的子域 L,而对应于子域 L 的是 G 中使得 L 的每个元素不变的置换组成的子群 H.如果 $H'\subset H$,则对应于 H' 的子域 L' 包含 L.例如,对于方程(3)的根域 E, E 的子域的下列图表(5)对应于伽罗瓦群的子群的图表(4).



[183]

(5)

D. 正规子群和单群

显然若尔当是第一位使用“同构”这个词的数学家.他对从群 G 到群 G' 上的我们现在称为的“同构”使用“全同构”这个词.若尔当还引进了另一个基本概念.对于满足下列条件的映射 f: $G \longrightarrow G'$,他使用“亚同构”这个词:

1) 对 $x \in G, y \in G$,有 $f(xy)=f(x)f(y)$;

2) 每个 $x' \in G'$ 至少对一个 $x \in G$ 具有 $f(x)$ 的形式.对于情形 2),我们现在称 f 是满射且写为

$$f(G)=G'.$$

只满足上述第一个条件的映射 f 现在称为从 G 到 G' 中的一个同态.此时有 $f(e)=e'$,e' 是 G' 的中性元,因为

$$f(x)=f(ex)=f(e)f(x).$$

由此得到

$$f(x^{-1})=(f(x))^{-1},$$

因为 $$e'=f(e)=f(xx^{-1})=f(x)f(x^{-1}).$$

由此得到下述结论:对于每个同态 f,$f(G)$ 是 G' 的子群.它称为 G 在 f 下的象.

一般地说,会有异于 e 的元素 $x \in G$,使得 $f(x)=e'$.这样的元素构成 G 的一个子群 N,称为 f 的核.事实上,如果 $f(x)=f(y)=e'$,则也有

$$f(xy)=f(x)f(y)=e', f(x^{-1})=(f(x))^{-1}=e'.$$

但这个子群 N 并不是一个任意的子群,它具有下述附加性质:对每个 $z \in G$ 和 $x \in N$,还有

$$zxz^{-1} \in N,$$

这是因为 $f(zxz^{-1})=f(z)f(x)f(z^{-1})=f(z)f(z^{-1})=e'$.

具有上述性质的子群称为正规子群(也称特异子群).这是伽罗瓦首次引进的.是他对 G 的任意子群 H,想到如同柯西所

做的那样(§2,B),以"模 H"的右等价类①安排 G 的元素.这些类是当 h 取遍 H 的所有元素时元素 sh 构成的形如 sH 的集合.如果两个类 sH 与 $s'H$ 具有公共元素 $sh = s'h'$,则有 $s' = shh'^{-1}$,从而对每个 $h'' \in H$,有

$$s'h'' = shh'^{-1}h'' \in sH.$$

换言之,$s'H \subset sH$;类似地有 $sH \subset s'H$,因此这两个类恒同.这样,G 是不同类的并,其中任何两个类没有公共元素.伽罗瓦以同样方式考虑左等价类 Hs.但一般 $Hs \neq sH$,只当 $sHs^{-1} = H$ 即 H 是正规子群时,才有对每个 $s \in G$,$Hs = sH$.

当 N 是正规子群时,我们能在等价类之间定义一个合成律,因为我们有 $(sN)(tN) = s(Nt)N = (st)N$,这是由于等式 $tN = Nt$,$NN = N$ 成立.易于验证这个合成律在等价类的集合(记[184]作 G/N)上定义了一个群结构,称为 G 关于 N 的商群.也易于验证映射 $f: s \longmapsto sN$ 是从 G 到 G/N 上的一个满同态,其核为 N.当 G 为有限时,G/N 的阶是 G 的阶除以 G/N 的阶之商,记作 $(G:N)$.

如果群 G 是交换的,则 G 的所有子群是正规的.例如,如果 G 是整数加法群 $\mathbf{Z}$ 而 N 是 $m(m>1)$ 的倍数构成的子群 $m\mathbf{Z}$,则 $\mathbf{Z}/m\mathbf{Z}$ 是模 m 的同余类的加法群(§2,C).

我们称群 G 是单的,如果它除 G 本身和仅含有中性元的子群 $\{e\}$ 外不包含任何别的正规子群.在有限交换群中,只有素数阶循环群是单的.事实上,设 G 是阶为 $n = rs(r>1, s>1)$ 的交换群.对 G 中每个 $x \neq e$,我们有 $x^n = e$,或 $(x^r)^s = e$.如果 $x^r = e$,则由 x 生成的子群不是 $\{e\}$,且它具有能整除 n 的阶,因而异于 G.如果 $y = x^r \neq e$,则有 $y^s = e$,从而有同样的结论.因此有限交换单群 G 是素数 p 阶的.此外,如果 $x \neq e$,则对 $1 \leqslant h \leqslant p-1$,

---

① 也称右陪集;下面说到的左等价类也称左陪集.两者相同时则称为陪集.——译注

所有幂 x^h 都互不相同,因为否则就会有 $x^{h-k}=e$, $1<h-k<p$,从而 $h-k$ 整除 p,这是不可能的.

若尔当和赫尔德证明,对每个有限群 G,存在互不相同的子群构成的序列

$$G=G_0\supset G_1\supset G_2\supset\cdots\supset G_{r-1}\supset G_r=\{e\}, \qquad (6)$$

使得 G_{k+1} 在 G_k 内是正规的,且每个群 G_k/G_{k+1} 是单的.

在例(4)中,三个 4 阶子群在 G 内是正规的和交换的,因而每个连接 G 和 $\{e\}$ 的子群链是若尔当 - 赫尔德序列(6). 当所有单群 G_k/G_{k+1} 为交换(因而为素数阶循环)群时,G 称为可解群. 可以证明,其阶只能被一个或两个素数整除或其阶不能被素数平方整除的群是可解的. 这种类型的定理中最值得注意的是费特 - 汤普森定理:每个奇数阶有限群是可解的. 其证明用反证法,长度竟超过 250 页! 是伽罗瓦引入了可解群概念,这给予他解决"用根式"解代数方程问题的钥匙:这样的解存在的必要充分条件是所给方程的伽罗瓦群是可解的.

E. 立方体的旋转

绕一直线 Δ 旋转(但非恒等旋转)的定义表明下述两个事实成立:如果点 M 关于此旋转不变,则它位于 Δ 上;如果 M 的变换 M' 异于 M,则通过 MM' 的中点且垂直于 MM' 的平面包含 Δ.

借助上述两个注记,易于确定所有使立方体 $K=ABCDA'B'C'D'$ 不变的旋转. 所给立方体的中心 O 不变,因此旋转的轴必通过 O. 如果顶点 A 不是不变的,则它或者变换到从 A 出发的一条边的端点,或者变换到位于通过 A 的一个面上的与 A 相对的顶点,最后,或者变换到 A 关于 O 的对称点 A'. 我们可用下述 [185] 事实来检验这四种情形:直线 OE 垂直于平面 $CDC'D'$,其中 E 是 AB 的中点(图 40).

容易看出我们要找的旋转有如下述(还有恒等旋转):

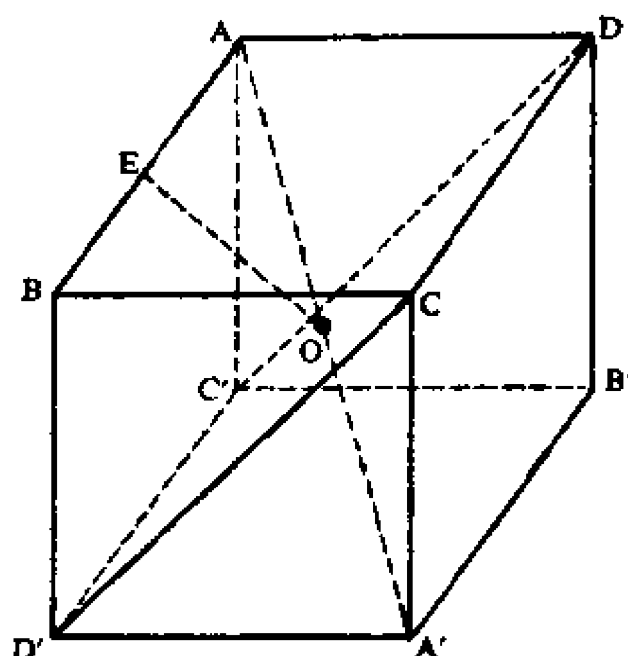


图 40

1) 绕连接 O 和一条边中点的连线旋转 180°;共有 6 个.每个这样的旋转连同恒等映射构成一个 2 阶循环子群.

2) 绕此立方体的对角线旋转 120°和 240°;对于每条对角线,这样的旋转连同恒等映射构成一个 3 阶循环子群.共有 8 个这样的旋转.

3) 绕垂直于一个面的直线旋转 90°,180°和 270°;对于每条这样的直线,所说的 3 个旋转连同恒等映射构成一个 4 阶循环子群.共有 9 个这样的旋转.

于是旋转总数为 $1+6+8+9=24$.

此立方体的旋转群 G 运算于从 O 出发垂直于立方体的面的 3 条垂线上,而且易见 G 以每种可能的方式对它们进行置换.如果每个旋转 $\sigma\in G$ 对应到这 3 条垂线相应的置换,则就定义了 G 到这 3 条直线的置换的群 $\mathfrak{S}_3$ 上的一个满同态.于是这个同态的核 N 是 G 的一个 4 阶正规子群,它由恒等映射和分别绕这 3 条直线转 180°的旋转组成.第二个 12 阶正规子群 $T\supset N$ 由 N 和绕所给立方体的 4 条对角线的旋转生成;它是使这 3 条

直线不变或对它们进行置换而不使任何一条固定的子群[①]. [186]

3. 关于环和域的补注

A. 模一素数的同余式

我们已看到(§2,C),对每个整数 $m>1$,模 m 同余类的加法 $(\bar{r},\bar{s})\longmapsto\bar{r}+\bar{s}$ 定义了这些类的集合上的一个 m 阶循环群结构.说 $\bar{r}+\bar{s}=\overline{r+s}$(由定义)意味着映射 $r\longmapsto\bar{r}$ 是整数加法群 $\mathbf{Z}$ 到所说同余类群上的一个满同态(附录2,D);此同态的核是 m 的倍数组成的子群 $m\mathbf{Z}$,因此模 m 同余类的加法群可记作 $\mathbf{Z}/m\mathbf{Z}$.但在 $\mathbf{Z}/m\mathbf{Z}$ 上还有第二个合成律,即 $(\bar{r},\bar{s})\longmapsto\bar{r}\cdot\bar{s}=\overline{rs}$(§1,F),它是结合的和交换的,并可直接验证有

$$\bar{r}\cdot(\bar{s}+\bar{t})=\bar{r}\cdot\bar{s}+\bar{r}\cdot\bar{t}.$$

这样这第二个合成律在 $\mathbf{Z}/m\mathbf{Z}$ 上定义了一个交换环结构并以 $\bar{1}$ 为单位元.现在假定取环 $\mathbf{Z}/p\mathbf{Z}$,p 是素数.于是环 $\mathbf{Z}/p\mathbf{Z}$ 是交换域,记作 $\mathbf{F}_p$.事实上,如果 $0<r<p$,则 p 与 r 的最大公因数是 1,从而有贝祖恒等式(参见第四章附录3)

$$mr+np=1,$$

m,n 是两个整数.由于在 $\mathbf{Z}/p\mathbf{Z}$ 中 $\bar{p}=\bar{0}$,由此得到 $\bar{m}\cdot\bar{r}=\bar{1}$.换言之,$\mathbf{Z}/p\mathbf{Z}$ 中每个类 $\bar{r}\neq\bar{0}$ 都有逆.

这样,$\mathbf{F}_p$ 中类 $\bar{r}\neq\bar{0}$ 的集合 $\mathbf{F}_p^*$ 是(如同所有域的情形一样)一个阶为 $p-1$ 的乘法群.于是对每个 $\bar{r}\neq\bar{0}$,有 $\bar{r}^{p-1}=\bar{1}$.用同余类语言,这可写为:对每个不是 p 的倍数的自然数 r,有

$$r^{p-1}\equiv 1(\bmod p).$$

这是数论中第一批定理之一,由费马证明并以费马"小"定理著

---

① 点 A,C,B',D' 是一个正四面体的顶点,而子群 T 使该四面体不变.G 的另外12个旋转把四面体 $ACB'D'$ 变换为它关于点 O 的对称象 $A'C'BD$.

称.

但还有更多结果：群 $\mathbf{F}_p^*$ 是循环的，这是欧拉猜想而首次由高斯(用不同的术语)证明的. 可以假定 $p\neq 2$；我们要用到 3 条引理.

Ⅰ) 设 n 是任意的自然数且 $n>1$. 对每个自然数 $k\leqslant n$，令 d 是 k 与 n 的最大公因数，则群 $\mathbf{Z}/n\mathbf{Z}$ 中类 $\bar{k}$ 的阶等于 n/d. 事实上，令 $n=n'd$，$k=k'd$. 如果 $ak=ak'd$ 被 $n=n'd$ 整除，则 n' 整除 ak'，又因 n' 与 k' 的最大公因数是 1，故 n' 整除 a. 特别是，使得 $\bar{k}$ 生成 $\mathbf{Z}/n\mathbf{Z}$ 的唯一自然数 $k\leqslant n$ 是与 n 互素的数，即与 n 的最大公因数是 1 的自然数. 这样的数的个数记作 $\varphi(n)$，称为 n 的欧拉函数.

Ⅱ) 关系"d 整除 n"记作 $d\mid n$. 对每个 $n>1$，我们有

$$\sum_{d\mid n}\varphi(n)=n. \tag{1}$$

事实上，对 n 的每个因数 d，如果 $n=n'd$，则类 $\overline{n'}$ 在群 [187] $\mathbf{Z}/n\mathbf{Z}$ 中的阶为 d，而由 Ⅰ)得知恰有 $\varphi(d)$ 个自然数 $k\leqslant n$，使得 $\bar{k}$ 的阶为 d. 由于对 $1\leqslant k\leqslant n$，$\bar{k}$ 的阶是 n 的因数，故得关系式(1). 此外还可看出，对 n 的每个因数 d，$\mathbf{Z}/n\mathbf{Z}$ 只有一个 d 阶子群.

反之，如果 n 阶有限群 G 满足：对 n 的每个因数 d，至多存在 d 个元素 x，使得 $x^d=e$，则 G 是循环的. 事实上，如果存在 d 阶元素 y，则子群 $\{e,y,y^2,\cdots,y^{d-1}\}$ 的所有元素满足 $x^d=e$，而由 Ⅰ)，恰存在 G 的 $\varphi(d)$ 个元素，其阶为 d. 如果存在 n 的一个因数 d，使得 G 中没有 d 阶元素，则 n 就会等于和 $\sum\varphi(d')$，其中 d' 取遍异于 d 的因数的集合. 但这与(1)矛盾. 特别是，G 至少存在一个 n 阶元素，由此 G 是循环的.

Ⅲ) 设 K 是任一交换域，令

$$P(t)=t^q+a_1t^{q-1}+\cdots+a_q$$

是系数取自 K 的 q 次多项式，于是在 K 中至多有 $P(t)=0$ 的 q 个根. 此点可由对于 q 的归纳推理看出：假定对某个 $c\in K$ 有

[188] $\mathbf{Z}[i]$,其元素称为高斯整数.为避免混淆,通常整数(正或负)称为有理整数.对每个高斯整数 z,$N(z)$是非负整数.高斯整数是平面上由通过两个轴上整数坐标点且平行于轴的平行线所构成的网格的"节点"(图 41).

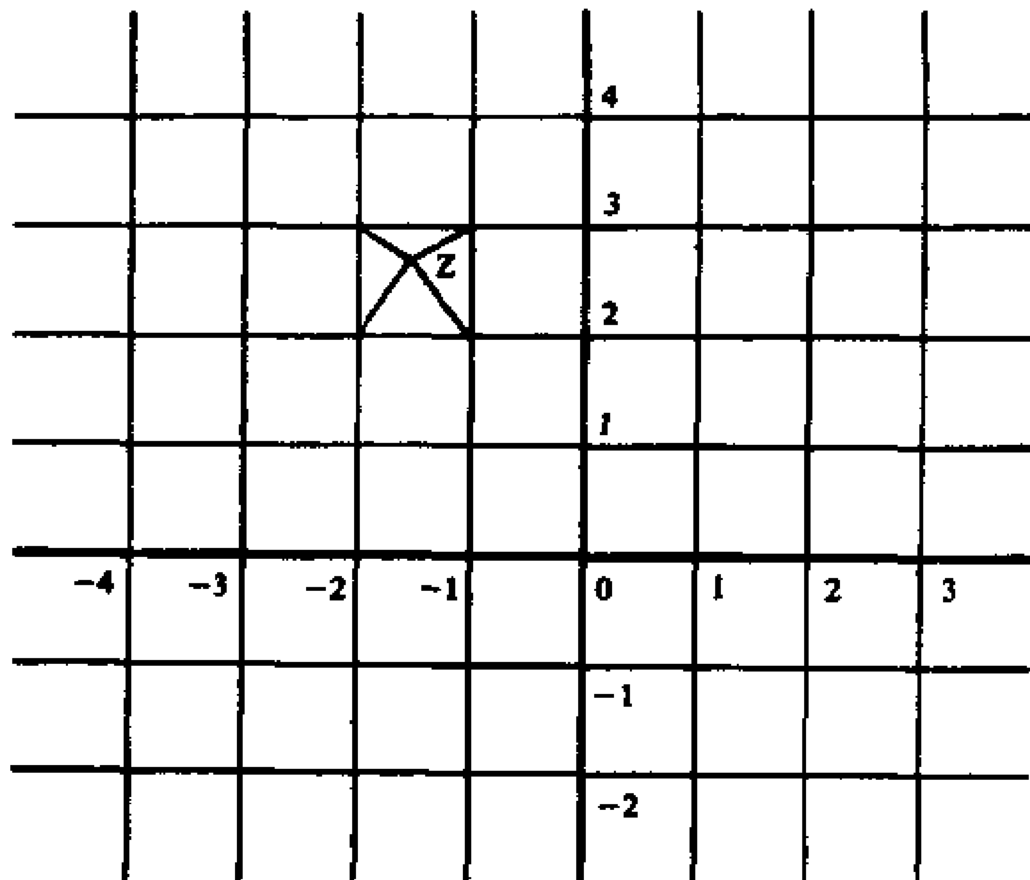


图 41

高斯整数的全部"算术"赖以作为基础的性质有如下述:对每个数 $z\in\mathbf{Q}(i)$,至少存在一个高斯整数 u,使得

$$N(z-u)\leqslant\frac{1}{2}.\tag{3}$$

这显然成立,因为 z 在上述网格的某一正方形(包括边)上,而此正方形的 4 个顶点中至少有一顶点到 z 的距离至多为半条对角线长 $1/\sqrt{2}$(图 41).

由此推出在环 $\mathbf{Z}[i]$ 中可有欧几里得除法：如果 u,v 是两个高斯整数，$v\neq 0$，则存在两个高斯整数 q,r，使得

$$u=qv+r,\quad N(r)\leqslant\frac{1}{2}N(v). \tag{4}$$

事实上，只须对 $\mathbf{Q}(i)$ 中的数 u/v 应用(3)，这表明存在高斯整数 q，使得

$$N\left(\frac{u}{v}-q\right)\leqslant\frac{1}{2}.$$

[189]

由此用(2)推出

$$N(u-qv)=N(v)N\left(\frac{u}{v}-q\right)\leqslant\frac{1}{2}N(v).$$

欧几里得算术的所有论证依赖于欧几里得除法和下述事实：如果非负整数的一个递减序列

$$r_1>r_2>\cdots>r_n>\cdots$$

只能终止于 0 这项，则此序列只有有限多项且最后一项为 0. 把这个注记应用于高斯整数的范数的序列

$$N(r_1)\geqslant 2N(r_2)\geqslant\cdots\geqslant 2^{n-1}N(r_n)\geqslant\cdots,$$

即见此序列只有有限项并终止于 $r_n=0$.

这样，有理整数的性质就能基本不变地移植到高斯整数上来. 为此只须简单地考虑到下述事实：在环 $\mathbf{Z}[i]$ 中，不仅数 1 和 -1，而且还有数 i 和 $-i$ 有 $\mathbf{Z}[i]$ 中的逆. 此外，只有这 4 个数具有这个性质，因为如果 z 有逆 z'，则有 $N(z)N(z')=1$，因为 $N(z)$ 是非负整数，由此得 $N(z)=1$. 我们称这 4 个整数 $\pm 1,\pm i$ 为 $\mathbf{Z}[i]$ 的单位. 按照定义，*素高斯整数*是异于 4 个单位的高斯整数 u，满足：如果 $u=vw$ 是两个高斯整数之积，则其中之一是单位. 对每个素高斯整数 u，4 个高斯整数 $\pm u,\pm iu$ 也是素的；我们称它们是*相伴的*. 其中只有一个 $u=a+ib$ 满足 $a>0,b\geqslant 0$；它称为*第一象限素高斯整数*. 如果 $u\in\mathbf{Z}[i]$ 是素的，则 $\bar{u}$ 也是素的；但与其共轭 $\bar{u}$ 相伴的第一象限非实素高斯整数只有 u

$=1+i$,因为 $1-i=-i(1+i)$.

高斯整数算术的基本结果是素整数分解定理:每个高斯整数 z 能以唯一方式写为

$$z=\epsilon u_1^{\alpha_1}u_2^{\alpha_2}\cdots u_k^{\alpha_k}, \tag{5}$$

其中 u_j 是互不相同的第一象限素高斯整数,α_j 是自然数,ϵ 是单位.

由此可得两个推论:

a) 如果一个素整数整除两个整数之积,则它整除其中之一.

b) 如果两个不相伴素整数整除同一整数,则其积整除该整数.

如果两个整数 z,z' 由其素整数分解给定,即(5)和

$$z'=\epsilon' v_1^{\beta_1}v_2^{\beta_2}\cdots v_h^{\beta_h}, \tag{6}$$

则 z,z' 的最大公因数可如下得到:取出现于分解式(5),(6)两者的素整数 $w_1,\cdots,w_l$,作乘积

$$d=w_1^{\gamma_1}w_2^{\gamma_2}\cdots w_l^{\gamma_l},$$

其中每个指数 γ_j 是出现于(5),(6)中 w_j 的两个指数中最小的数.我们有贝祖恒等式

$$d=az+bz', \tag{7}$$

其中 a,b 是 $\mathbf{Z}[i]$中的两个整数.其证明与 $\mathbf{Z}$ 中基于欧几里得
[190] 除法证明贝祖恒等式(见第四章,附录3)相同.

$\mathbf{Z}$ 中的素数与 $\mathbf{Z}[i]$中的素数之间的关系可完整陈述如下.

Ⅰ) 素数2是(不计相异一个单位)一个素高斯整数的平方,因为

$$2=(1+i)(1-i)=-i(1+i)^2.$$

Ⅱ) 如果素有理整数 $p\neq 2$ 不是 $\mathbf{Z}[i]$中的素数,则它被一个第一象限非实素整数 u 整除,因而也被其共轭 $\bar{u}$ 整除.由于 $p\neq 2$,故 u 与 $\bar{u}$ 不是相伴的,因而 p 被 $u\bar{u}=N(u)$整除.但因 p 是 $\mathbf{Z}$ 中的素数,故必有 $p=N(u)$.如果令 $u=x+iy$, $x>0$, $y>0$,则有 $p=x^2+y^2$,这蕴涵 p 具有 $4k+1$ 的形式.

Ⅲ）如果 w 在 $\mathbf{Z}[i]$中是素的，我们如同 A)定义 $\mathbf{Z}[i]$中模 w 的同余；由贝祖恒等式(7)得知环 $\mathbf{Z}[i]/w\mathbf{Z}[i]$中每个不等于 $\bar{0}$ 的类都有逆．换言之，$\mathbf{Z}[i]/w\mathbf{Z}[i]$是域．如果 p 是形如 $4k+1$ 的有理素数，我们在 A)中已看到存在有理整数 x，使得 $x^2+1\equiv 0 \pmod p$．这意味着在 $\mathbf{Z}[i]$中有

$$(x+i)(x-i)\equiv 0 \pmod p,$$

又由于我们也已知道 $x+i$ 和 $x-i$ 在 $\mathbf{Z}[i]$中都不能被 p 整除，所以 $x+i$ 和 $x-i$ 的类在 $\mathbf{Z}[i]/p\mathbf{Z}[i]$中是零因子，因之 $\mathbf{Z}[i]/p\mathbf{Z}[i]$不是域，从而 p 在 $\mathbf{Z}[i]$中不是素数．现取第一象限素整数 u，使 u 整除 p．而 u 不可能属于 $\mathbf{Z}$，故由Ⅱ)得到 $p=N(u)$，从而 u 是整除 p 的唯一的第一象限素整数．如果令 $u=a+ib$，则有 $p=a^2+b^2$，因而每个形如$4k+1$的有理素数以唯一的方式表示为两个自然数平方之和；这就是费马－欧拉定理(见第四章，附录3)．

当欧几里得除法能以同样方式推广到环 $\mathbf{Z}[\sqrt{D}]$中时，就能推导出完全类似的定理．可惜已经证明，这只对有限多个不能被一平方整除的整数 D 成立．对于 $D<0$，D 的可能的值是 -1，-3，-7，-11．事实上，存在无穷多个 D 值，对于它们素因数分解不是唯一的．例如，对于 $D=-5$，我们有

$$9=3\cdot 3=(2+\sqrt{-5})(2-\sqrt{-5}),$$

且可证明 $3, 2+\sqrt{-5}, 2-\sqrt{-5}$在 $\mathbf{Z}[\sqrt{-5}]$中是素的．库默尔、克罗内克和戴德金发现，通过引进新的对象——理想和赋值，有可能重建“唯一分解”定理；随同这一发现，代数数论也由之诞生．

C．模一多项式的同余式

本节要给出“同余式演算”的第三个例子．设 K 是具有无穷多个元素的域．系数在 K 中的多项式是映射

$$x \longmapsto P(x)=a_0x^n+\cdots+a_n, \tag{8}$$

它是有限个单项式 $a_j x^{n-j}$之和．称多项式 P 是 n 次的（写为 $\deg(p)=n$），如果 P 的最高次单项式a_0x^n 具有非零系数．零次 [191]

多项式是非零常数 $c\in K$. 显然系数在 K 中的多项式关于多项式加法和乘法构成一个交换环,记作 $K[x]$. 由于 K 有无穷多个元素,所以多项式(8)只当 $a_0=a_1=\cdots=a_n=0$ 时才能恒等于零,因为在 A)中我们已看到 n 次多项式不可能在多于 K 的 n 个元素处取零值. 注意 m 次多项式与 n 次多项式之积

$$(a_0x^m+\cdots+a_m)(b_0x^n+\cdots+b_n)$$

为 $m+n$ 次,因为在这个积中只有一个 $m+n$ 次单项式,即 $a_0b_0x^{m+n}$,而因 K 是域,故当 $a_0\neq 0$, $b_0\neq 0$ 时 $a_0b_0\neq 0$.

由此推出,$K[x]$ 中具有逆的仅有元素是常数 $c\neq 0$.

这个环的基本性质是在其上也能定义欧几里得除法:如果 A,B 是两个多项式,$B\neq 0$,则存在两个多项式 Q,R,使得

$$A=BQ+R,\text{而 } R=0 \text{ 或 } \deg(R)<\deg(B). \tag{9}$$

事实上,设 m,n 分别是 A,B 的次数,如果 $A=0$ 或 $m<n$,则只须取 $Q=0$, $R=A$ 即可得(9). 如果不是上述情形,则能求出单项式 $M_1=cx^{m-n}$,满足:

如果 $A_1=A-BM_1$,则有

$$A_1=0 \text{ 或 } \deg(A_1)<\deg(A). \tag{10}$$

事实上,如果 $A=a_0x^m+\cdots+a_m$, $B=b_0x^n+\cdots+b_n$,则只须取 $c=a_0/b_0$ 即得(10).

重复上述过程,可得多项式序列:

$$A_1=A-BM_1,\text{而 } A_1=0 \text{ 或 } \deg(A_1)<\deg(A),$$

$$A_2=A_1-BM_2,\text{而 } A_2=0 \text{ 或 } \deg(A_2)<\deg(A_1),$$

$$\cdots\cdots$$

在至多等于 $m-n+1$ 的有限步后,必定达到

$$A_{k+1}=A_k-BM_{k+1},\text{而 } A_{k+1}=0 \text{ 或 } \deg(A_{k+1})<n.$$

于是我们有

$$A_{k+1}=A-B(M_1+M_2+\cdots+M_{k+1}),$$

令 $Q=M_1+M_2+\cdots+M_{k+1}$, $R=A_{k+1}$,上式即为(9).

满足(9)的多项式 Q,R 是唯一确定的.事实上,如果还有 $A=BQ_1+R_1$,而 $R_1=0$ 或 $\deg(R_1)<\deg(B)$,则就会有

$$B(Q-Q_1)=R_1-R,$$

如果 $R_1-R\neq 0$,则有 $\deg(R_1-R)<\deg(B)$,而在 $Q-Q_1\neq 0$ 情形下,这与上式矛盾,因此必有 $Q-Q_1=0$, $R-R_1=0$.我们分别称 Q,R 为 A 除以 B 的欧几里得除法的商式和余式.

这样,全部欧几里得算术论证都可直接移植过来.此地对应于素数的是不可约多项式,即满足下述条件的非常数多项式 P:如果 $P=AB$ 是两个多项式之积,则 A,B 中必有一为常数.如果 P 是不可约的,则对每个非零的 $c\in K$, cP 也是不可约的.在 [192] 这些多项式中,我们选取首一的,即其最高次单项式系数等于 1 的多项式.唯一分解为素因子的定理于是就如下述:$K[x]$ 中每个非常数多项式 $A\neq 0$ 能以唯一方式写为

$$A=cP_1^{\alpha_1}P_2^{\alpha_2}\cdots P_k^{\alpha_k}, \tag{11}$$

其中 $c\in K$, P_j 是互不相同的首一不可约多项式,α_j 是自然数.如同 B)中那样,我们能对两个非零多项式 A,B 定义最大公因式(在相差一个常数因子范围内),并推出贝祖恒等式

$$D=MA+NB, \tag{12}$$

其中 D 是 A,B 的最大公因式,M,N 是两个多项式.

接着可定义模一个非常数多项式 M 的同余式 $A\equiv B(\mathrm{mod}M)$;模 M 同余于多项式 A 的多项式类 $\bar{A}$ 是多项式 $A+QM$ 的集合,其中 Q 是任一多项式.这些类构成一个环 $K[x]/MK[x]$.欧几里得除法表明每个类 $\bar{A}$ 有唯一的一个"代表",它或是 0 或是 A 除以 M 的欧几里得除法的余式,该余式是次数小于 $\deg(M)$ 的多项式.贝祖恒等式(12)还表明,如果 P 是不可约多项式,则环 $K[x]/PK[x]$ 是域.

D. 代数函数域

给定无穷域 K 上的两个多项式 $x\longmapsto P(x)$, $x\longmapsto$

$Q(x)$, Q 不恒等于零,"分式" P/Q 的"自然"定义当是函数

$$x \longmapsto P(x)/Q(x).$$

但此函数在使得 $Q(x)=0$ 的点 $x \in K$ 处没有定义.这样的点只有有限多个,但它们依赖于 Q;这些点的出现是在"分式"上定义运算困难的原因所在.例如,函数

$$x \longmapsto \frac{1}{x(x-1)}, x \longmapsto \frac{1}{x}$$

对 $x=0$ 没有定义,但因对于 $x \neq 1$,有

$$\frac{1}{x(x-1)}+\frac{1}{x}=\frac{1}{x-1},$$

故所给两个函数之和对 $x=0$ 有定义.于是我们就不喜欢对"有理分式" P/Q 谈论它的"值",而喜欢考虑 $K[x]$ 的多项式偶 (A,B) 的类,其中 $B \neq 0$.两个偶 (A,B), (A_1,B_1) 称为属于相同的类,如果

$$AB_1 - BA_1 = 0.$$

然后有必要定义两个类的和与积.对于两个偶 (A,B), (C,D),考虑偶 $(AD+BC, BD)$, (AC, BD).必须证明,如果例如用同一类中的偶 (A_1,B_1) 代替 (A,B),"和"偶与"积"偶的类不改变.这 [193] 是恒等式

$$(AD+BC)B_1D-(A_1D+B_1C)BD=(AB_1-A_1B)D^2=0,$$

$$ACB_1D - BDA_1C=(AB_1-A_1B)CD=0$$

的推论.当用语随便时,我们写 A/B 来代替偶 (A,B) 的类,而对同一类的两个偶 (A,B), (A_1,B_1),有 $A/B=A_1/B_1$.可以直接验证这样定义的类构成一个域;当 $A \neq 0$ 时,类 A/B 的"逆"是 B/A.以 $K(x)$ 记这个域,称为系数在 K 中的单变量有理分式域.

如果 A 与 B 具有最大公因式 D,从而 $A=DA_1$, $B=DB_1$,则有 $A/B=A_1/B_1$,从而有理分式总能写为两个"互素的"多项式(即它们的最大公因式为 1)之商.但两个这种形式的有理分式

之和不一定仍是这种形式.

现设 K 是复数域 $\mathbf{C}$. 一条代数平面曲线是 $\mathbf{C}^2$ 的满足方程

$$a_0(x)y^n + a_1(x)y^{n-1} + \cdots + a_n(x) = 0 \qquad (13)$$

的点 (x, y) 的集合 Γ, 其中 $a_j(x)$ 是 $\mathbf{C}[x]$ 中的多项式, 且 $a_0(x) \neq 0$. 曲线 Γ 称为不可约的, 如果系数在域 $\mathbf{C}(x)$ 中的多项式

$$P(y) = y^n + \frac{a_1(x)}{a_0(x)}y^{n-1} + \cdots + \frac{a_n(x)}{a_0(x)}$$

是不可约的. 此时模 P 的同余类环 $\mathbf{C}(x)[y]/P\cdot\mathbf{C}(x)[y]$ 是一个域 L, 称为 Γ 上的有理函数域(或 Γ 上的代数函数域). 对这一名称可作如下解释: L 的每个元素可唯一地表示为

$$F = r_0(x)y^{n-1} + r_1(x)y^{n-2} + \cdots + r_{n-1}(x),$$

其中 $r_j(x)$ 是域 $\mathbf{C}(x)$ 的任意元素. (这些"代表"是环 $\mathbf{C}(x)[y]$ 中的多项式除以 P 的欧几里得除法的余式.)

设 (u, v) 是 $\mathbf{C}^2$ 的位于 Γ 上的点即满足方程(13). 如果 u 不是有理分式 $r_0, r_1, \cdots, r_{n-1}$ 的分母的零点, 则在 F 中可用 u 代替 x, 用 v 代替 y. 这样得到的复数称为 F 在点 (u, v) 处的值. 有可能在某种意义上"扩充" F 的值的定义到 u 是 r_j 的分母的零点的点处以及 Γ 的所谓"无穷远点"处, 使之得到一个真正的函数. 但这需要涉及称为"黎曼曲面"的较高深概念.

E. 关于有序域的注记

我们要在本小节中表明, 实系数有理分式域 $\mathbf{R}(x)$ 可赋予一个有序域(§3,F)结构. 首先对多项式

$$P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \qquad (14)$$

定义关系 $P \geqslant 0$; 它意味着或者 $P = 0$ (即 $a_0 = a_1 = \cdots = a_n = 0$), [194] 或者 $a_0 > 0$. 如果 P, Q 是两个多项式, 则定义 $P \leqslant Q$ 为在上述意义下 $Q - P \geqslant 0$. 最后, 对有理分式 $R = P/Q$, $Q \neq 0$, 定义关系 $R \geqslant 0$ 为在上述意义下 $PQ \geqslant 0$, 而对两个有理分式 R, S, 关系 $R \leqslant S$ 意味着 $S - R \geqslant 0$.

我们必须验证序结构公理(§3,D)和有序域结构公理(§3,F).下面进行三项验证.

1) 全序公理.为此只须证明,如果 $R\geqslant 0$, $-R\geqslant 0$,则 $R=0$,而由上面的定义,只须考虑多项式情形.但如果 $P(x)$是由(14)表示的多项式且 $P\geqslant 0$, $-P\geqslant 0$,则不可能有 $P\neq 0$,因为这蕴涵对于实数 a_0 有 $a_0>0$ 和 $a_0<0$,这是不可能的.

2) 如果 $R\geqslant 0$, $S\geqslant 0$,则 $RS\geqslant 0$.由定义又可限于验证多项式情形.但如果 $P\geqslant 0$, $Q\geqslant 0$, P 是多项式(14), Q 是多项式

$$Q(x)=b_0x^m+\cdots+b_m, \tag{15}$$

则或者关系 $P=0$, $Q=0$ 有一成立从而 $PQ=0$;或者有 $a_0>0$, $b_0>0$,而 PQ 中最高次项是$a_0b_0x^{m+n}$,且 $a_0b_0>0$.同样论证表明,对每个多项式 P 有 $P^2\geqslant 0$.

3) 如果 $R\geqslant 0$, $S\geqslant 0$,则 $R+S\geqslant 0$.设 $R=A/B$, $S=C/D$,则 $R+S=(AD+BC)/BD$.问题归结为证明,如果对 4 个多项式 A,B,C,D 有$AB\geqslant 0$, $CD\geqslant 0$,则有

$$(AD+BC)BD=ABD^2+CDB^2\geqslant 0.$$

由 2),这导致证明,如果多项式(14),(15)都$\geqslant 0$,则 $P+Q\geqslant 0$.当 $P=0$ 或 $Q=0$ 时,这显然成立.如果不是这种情形,设例如 $m>n$,则 $P+Q$ 中最高次项是b_0x^m,而 $b_0>0$.现在只剩下考虑 $m=n$ 的情形,此时 $P+Q$ 中最高次项是$(a_0+b_0)x^m$,而由于 $a_0>0$, $b_0>0$,故 $a_0+b_0>0$.

这个例子的意义在于阿基米德公理(第三章§6及附录2)不满足.事实上,实数是0次多项式,而对两个实数 r,s,关系 $r\leqslant s$ 在 $\mathbf{R}$ 和 $\mathbf{R}(x)$中是相同的.但在 $\mathbf{R}(x)$中,对每个实数 r,多项式 x 大于r;可以说 x 关于实数是"无穷大";又由于对每个实数 $r>0$,有

$$0<\frac{1}{x}<\frac{1}{r},$$

于是可以说有理分式 $1/x$ 在 $\mathbf{R}(x)$中是"无穷小".

我们方才用到这样的事实：在 $\mathbf{R}(x)$ 中，任一多项式的平方 $\geqslant 0$. 这是一个一般的事实：对每个有序域 K 和每个元素 $x \in K$，有 $x^2 \geqslant 0$. 事实上，如果 $x \geqslant 0$，则从有序域公理直接得到 $x^2 \geqslant 0$；如果 $x < 0$，则 $-x > 0$，从而 $x^2 = (-x)^2 > 0$. 特别是，这就证明复数域 $\mathbf{C}$ 上没有有序域结构，因为 [195]

$$i^2 = -1 = -(1)^2,$$

因而如果 $\mathbf{C}$ 上有有序域结构，就会同时有 $i^2 > 0$ 和 $i^2 < 0$，这是荒谬的.

更一般地说，在有序域 K 中，关系

$$x_1^2 + x_2^2 + \cdots + x_m^2 = 0 \tag{16}$$

只当 $x_1 = x_2 = \cdots = x_m = 0$ 时才能成立. 事实上，假定 $x_1 \neq 0$，从而 $x_1^2 > 0$. 由于 $x_2^2 + \cdots + x_m^2$ 是正元素之和，故 $x_2^2 + \cdots + x_m^2 \geqslant 0$；于是有

$$0 < x_1^2 \leqslant x_1^2 + (x_2^2 + \cdots + x_m^2),$$

从而不可能出现(16). 例如，在有限域 $\mathbf{F}_p$(附录 3，A)上没有有序域结构，因为在该域中 $\bar{1} = (\bar{1})^2$ 且

$$\bar{0} = \bar{p} = \bar{1} + \bar{1} + \cdots + \bar{1} (p \text{ 项}).$$

阿廷和施赖埃尔证明，在一个交换域 K 中，关系式(16)对非零元素不可能成立对于 K 上有序域结构的存在性不仅是必要的，而且还是充分的.

4. 距离的例子

A. 连续函数空间中的距离

分析中最有用的"函数空间"之一是定义在区间 $0 \leqslant x \leqslant 1$ 上的实值连续函数的空间 $C[0,1]$. 定义在该空间上记作 $d_\infty(f,g)$ 的距离(§3，E)由下述公式给出：对于 $C[0,1]$ 中的两个函数 f, g，

$$d_\infty(f,g) = \max\{|f(x) - g(x)| ; 0 \leqslant x \leqslant 1\}, \tag{1}$$

这里式(1)的右边是使得 $|f(x)-g(x)|\leqslant\alpha$ 对所有满足 $0\leqslant x\leqslant1$ 的 x 成立的最小实数 $\alpha\geqslant0$. 如果 f,g,h 是 $C[0,1]$ 中的 3 个函数,则对所有 x,

$$f(x)-h(x)=(f(x)-g(x))+(g(x)-h(x)),$$

从而

$$\begin{aligned}|f(x)-h(x)|&\leqslant|f(x)-g(x)|+|g(x)-h(x)|\\&\leqslant d_\infty(f,g)+d_\infty(g,h),\end{aligned}\tag{2}$$

于是由定义得到

$$d_\infty(f,h)\leqslant d_\infty(f,g)+d_\infty(g,h).\tag{3}$$

这是关于距离的公理(iii)("三角形不等式")(§3,E);公理(i),(ii)显然满足.

关于该距离的"邻域"有简单的几何解释:说 $d_\infty(f,g)\leqslant\alpha$ 意味着 g 的图象包含在满足下述条件的点 (x,y) 形成的宽度为 2α 的"带域"中:

$$0\leqslant x\leqslant1,f(x)-\alpha\leqslant y\leqslant f(x)+\alpha.$$

[196] (见图 42.)

关于该距离的"极限"概念在分析中十分重要. $C[0,1]$ 中函数的一个序列 (f_n) 称为一致收敛到函数 f, 如果数列 $(d_\infty(f,f_n))$ 收敛到 0.

上述这些定义对于定义在 $0\leqslant x\leqslant1$ 上且只假定为有界(不一定连续)的函数也有效. 我们有一条基本定理:如果连续函数序列 (f_n) 一致收敛到函数 f(它必定有界),则 f 必是连续的.

在 $C[0,1]$ 上还能定义另一个记作 $d_1(f,g)$ 的距离,它由公式

$$d_1(f,g)=\int_0^1|f(x)-g(x)|\mathrm{d}x\tag{4}$$

给出. 这里三角形不等式

$$d_1(f,h)\leqslant d_1(f,g)+d_1(g,h)$$

仍由对于 $[0,1]$ 上所有 x 成立的不等式(2)和下述事实得出:如果对于 $C[0,1]$ 中的函数 u,v 有 $u(x)\leqslant v(x)$,则

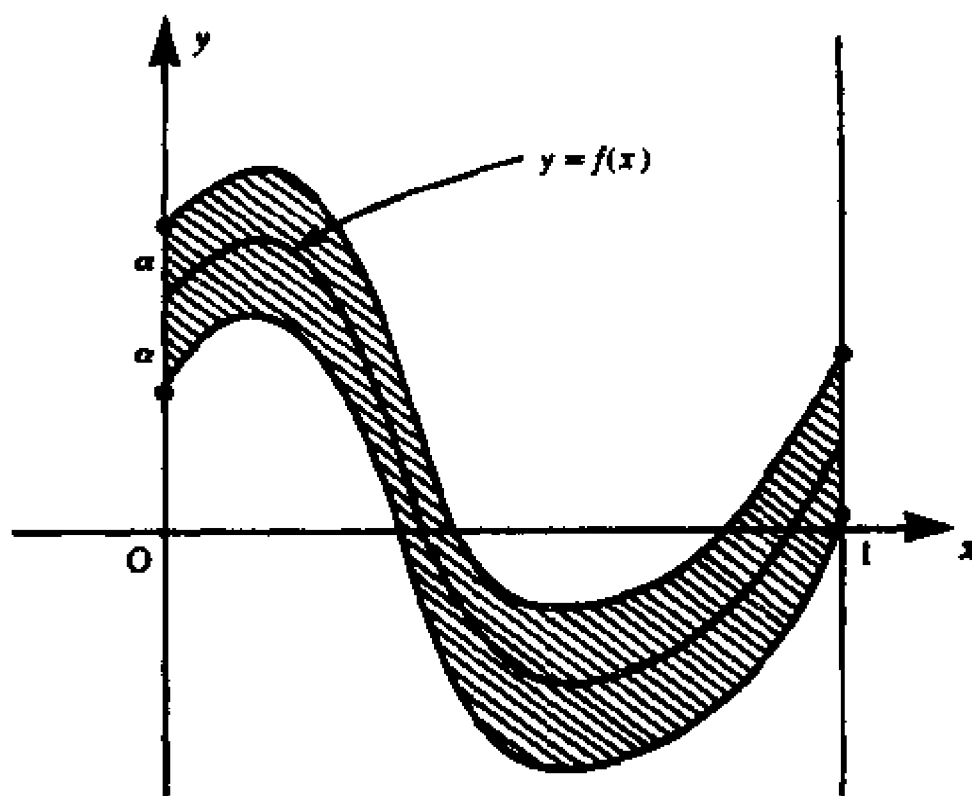


图 42

$$\int_0^1 u(x)\mathrm{d}x \leqslant \int_0^1 v(x)\mathrm{d}x.$$

只有关于距离的公理(i)需要证明:如果 $d_1(f,g)=0$,则 $f=g$.
事实上,令

$$\phi(x)=|f(x)-g(x)|,$$

已知

$$F(x)=\int_0^x \phi(t)\mathrm{d}t$$

的导数在每个点处等于 $\phi(x)$(第三章,§9).由于对[0,1]中每个 x 有 $F'(x)\geqslant 0$,所以函数 F 或是递增或是常数.但由于 [197] $F(0)=0$且由假定 $F(1)=d_1(f,g)=0$,我们必定有 $F(x)=0$ 对于 $0\leqslant x\leqslant 1$ 成立,因此 $\phi(x)=F'(x)=0$,即 $f=g$.

关于距离 d_1 的"邻域"概念也有简单的几何解释:说 $d_1(f,g)\leqslant a$ 意味着包含于 f 与 g 的图象"之间"的面积至多为 a(见图 43).

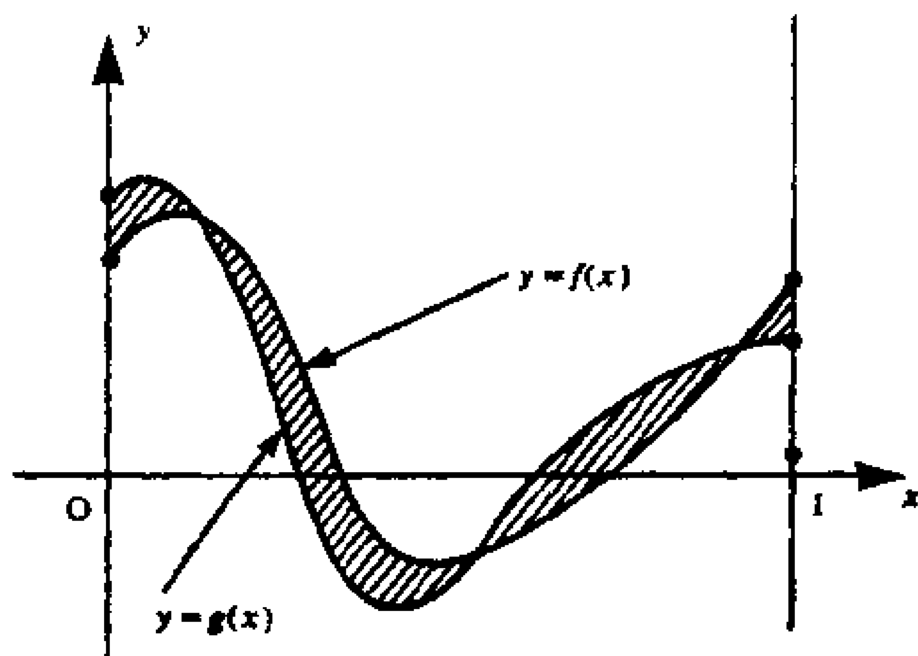


图 43

显然我们有

$$d_1(f,g) \leqslant d_\infty(f,g). \tag{5}$$

于是,如果连续函数序列(f_n)一致收敛到f,则也有

$$\lim_{n\to\infty}\int_0^1 f_n(x)\mathrm{d}x = \int_0^1 f(x)\mathrm{d}x,$$

因为

$$\left|\int_0^1 f(x)\mathrm{d}x - \int_0^1 f_n(x)\mathrm{d}x\right| = \left|\int_0^1 (f(x) - f_n(x))\mathrm{d}x\right| \leqslant d_1(f,f_n).$$

另一方面,可能发生这样的情形:序列(f_n)关于距离d_1收敛到f,但关于距离d_∞不收敛到f.为此取

$$f_n(x) = \begin{cases} 2nx & \left(0 \leqslant x \leqslant \dfrac{1}{2n}\right), \\ 2 - 2nx & \left(\dfrac{1}{2n} \leqslant x \leqslant \dfrac{1}{n}\right), \\ 0 & \left(\dfrac{1}{n} \leqslant x \leqslant 1\right). \end{cases}$$

我们有

$$\int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{2n}.$$

因而$(d_1(f_n,0))$收敛到0,然而对所有 n 有$d_\infty(f_n,0)=1$.(见图44.)

这样我们就在同一个函数集 $C[0,1]$上定义了两个不等价的距离. [198]

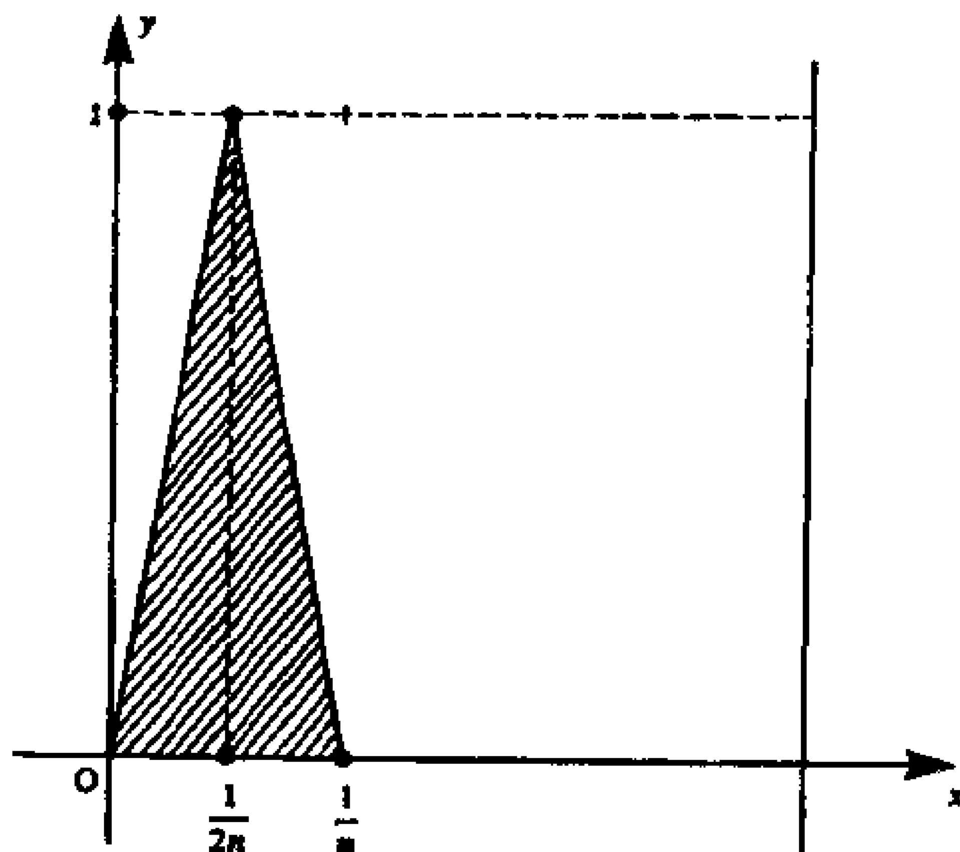


图 44

B. 准希尔伯特空间

设 E 是域 $\mathbf{R}$ 上的向量空间(§3,F).推广通常空间中标量积概念(§1,B),让我们假定给定了从 $E\times E$ 到 $\mathbf{R}$ 中具有下列性质的映射

$$(x,y)\longmapsto(x\mid y):$$

(I) $(x\mid y)=(y\mid x)$;

(ii) $(x+x'|y)=(x|y)+(x'|y)$;

(iii) 对于所有 $\lambda\in\mathbf{R}$,$(\lambda x|y)=\lambda(x|y)$;

(iv) 除 $x=0$ 外,$(x|x)>0$.

这样的映射仍称为 E 上的标量积,而在向量空间 E 上定义了标量积的结构称为准希尔伯特空间结构.

直接推广通常空间 $\mathbf{R}^3$ 的例子是向量空间 $\mathbf{R}^n$,这里 n 是任一大于 3 的自然数,而对两个向量 x,y,令

$$(x|y)=\xi_1\eta_1+\xi_2\eta_2+\cdots+\xi_n\eta_n,$$

其中 ξ_j 是 x 的分量,η_j 是 y 的分量(§1,B).易于直接验证公理(i)到(iv).另一例子是区间[0,1]上实值连续函数的空间 $C[0,1]$(上一小节);这里标量积由

[199]

$$(f|g)=\int_0^1 f(x)g(x)\mathrm{d}x \tag{6}$$

定义.公理(i),(ii),(iii)的验证是显然的.至于(iv),必须证明,如果

$$(f|f)=\int_0^1 f^2(x)\mathrm{d}x=0,$$

则对所有 $x\in[0,1]$必有$f(x)=0$.这可通过考虑原函数

$$F(x)=\int_0^x f^2(t)\mathrm{d}t,$$

用类似于 A)中的论证来进行.

标量积的基本性质是柯西-布尼亚科夫斯基-施瓦茨不等式

$$|(x|y)|\leqslant\sqrt{(x|x)}\cdot\sqrt{(y|y)}, \tag{7}$$

等式仅当对于两个不全为零的标量 α,β 有 $\alpha x+\beta y=0$ 时才出现.

事实上,注意对每个实数 t,有

$$(x+ty|x+ty)\geqslant 0,$$

它给出

$$(x|x)+2t(x|y)+t^2(y|y)\geqslant 0. \tag{8}$$

如果$(y|y)=0$(即$y=0$),则也有$(x|y)=0$.否则,在(8)中使t取值

$$t_0=-\frac{(x|y)}{(y|y)},$$

得到

$$(x|x)-\frac{(x|y)^2}{(y|y)}\geqslant 0, \tag{9}$$

即

$$(x|y)^2\leqslant(x|x)(y|y),$$

由此取正平方根即得(7).此外,如果(9)的左边为零,则它意味着

$$(x+t_0y|x+t_0y)=0,$$

就是说$x+t_0y=0$.

令$\|x\|=\sqrt{(x|x)}$,不等式(7)可写为

$$|(x|y)|\leqslant\|x\|\cdot\|y\|. \tag{10}$$

由上述不等式可得闵科夫斯基不等式:

$$\|x+y\|\leqslant\|x\|+\|y\|. \tag{11}$$

事实上,我们有

$$\begin{aligned}\|x+y\|^2&=(x+y|x+y)\\&=(x|x)+2(x|y)+(y|y)\\&\leqslant(x|x)+2|(x|y)|+(y|y)\\&\leqslant\|x\|^2+2\|x\|\cdot\|y\|+\|y\|^2\quad(\text{由}(10))\\&=(\|x\|+\|y\|)^2,\end{aligned}$$

取平方根即得(11). [200]

闵科夫斯基不等式使我们能通过

$$d(x,y)=\|x-y\| \tag{12}$$

在准希尔伯特空间E上定义一个距离,因为由(iv)可知仅当$x=y$时才能有$d(x,y)=0$.由(i)和(iii)推出$(-z|-z)=(z|z)$,从而$d(y,x)=d(x,y)$.最后,由闵科夫斯基不等式,有

$$\|x-z\|=\|(x-y)+(y-z)\|\leqslant\|x-y\|+\|y-z\|,$$

即 $d(x,z)\leqslant d(x,y)+d(y,z)$. 当然,由此我们重新发现 $\mathbf{R}^3$ 中的欧几里得三角形不等式.

我们看到,我们也在 $C[0,1]$上定义了第三个距离

$$d_2(f,g)=\sqrt{\int_0^1 (f(x)-g(x))^2\mathrm{d}x}. \tag{13}$$

容易证明此距离既不等价于距离 $d_\infty(f,g)$,也不等价于距离 $d_1(f,g)$.

C. 希尔伯特空间

我们以 l^2 或 $l^2_{\mathbf{R}}$ 记满足下述条件的实数无穷序列 $x=(x_n)_{n=1,2,\cdots}$的集合:非负项级数

$$\sum_{n=1}^{\infty} x_n^2 = x_1^2+x_2^2+\cdots+x_n^2+\cdots$$

收敛. 首先在此集合上定义一个域 $\mathbf{R}$ 上的向量空间结构(§3,F):l^2 的两个序列 $x=(x_n)$与 $y=(y_n)$的和定义为序列

$$x+y=(x_n+y_n);$$

必须证明级数

$$\sum_{n=1}^{\infty}(x_n+y_n)^2$$

收敛,但这是不等式

$$(x_n+y_n)^2\leqslant 2(x_n^2+y_n^2)$$

的直接推论. 标量 $\lambda\in\mathbf{R}$ 乘 $x=(x_n)$定义为

$$\lambda x=(\lambda x_n),$$

而由$(\lambda x_n)^2=\lambda^2x_n^2$,显见级数

$$\sum_{n=1}^{\infty}(\lambda x_n)^2$$

收敛,从而 $\lambda x\in l^2$. 向量空间的公理(§3,F)可以直接验证.

在此向量空间上由

$$(x|y)=\sum_{n=1}^{\infty}x_ny_n \tag{14}$$

定义一个标量积. 显然必须证明上式右边的级数收敛. 事实上,

作为不等式

$$|x_n y_n| \leqslant \frac{1}{2}(x_n^2 + y_n^2)$$

的推论，它是绝对收敛的. B)中公理(i),(ii),(iii)的验证是显然的；[201]
至于(iv)，为证明若$(x|x)=0$即

$$x_1^2 + x_2^2 + \cdots + x_n^2 + \cdots = 0,$$

则每个x_j为零，只须注意对每个j有

$$0 \leqslant x_j^2 \leqslant x_1^2 + x_2^2 + \cdots + x_n^2 + \cdots.$$

赋予标量积(14)的空间$l_{\mathbf{R}}^2$称为希尔伯特空间(或实希尔伯特空间).

D. p 进 距 离

设p是素数.每个非零整数m能写成

$$m = p^{\alpha} m',$$

其中整数m'不被p整除，α是非负整数.数α称为m的p进赋值，记为$v_p(m)$.于是我们有

$$v_p(m) = 0, \text{如果 } p \text{ 不整除 } m; \tag{15}$$

$$v_p(-m) = v_p(m); \tag{16}$$

$$v_p(mn) = v_p(m) + v_p(n). \tag{17}$$

最后，我们有下述不等式：如果$m+n \neq 0$，则

$$v_p(m+n) \geqslant \min\{v_p(m), v_p(n)\}, \tag{18}$$

右边表示两个数$v_p(m)$，$v_p(n)$中较小的数.为证明(18)，写$m = p^{\alpha} m'$，$n = p^{\beta} n'$，p既不整除m'，也不整除n'.假定例如$\alpha \geqslant \beta$，则

$$m + n = p^{\beta}(p^{\alpha-\beta} m' + n').$$

由于在$\alpha=\beta$的情形可能发生p整除$p^{\alpha-\beta}m'+n'$，故由上只能断言$v_p(m+n) \geqslant \beta$.然而，如果$\alpha > \beta$，则有等式

$$v_p(m+n) = \min\{v_p(m), v_p(n)\}.$$

现在来对每个整数 m 定义 p 进绝对值 $|m|_p$:如果 $m=0$,则令 $|0|_p=0$. 否则,令

$$|m|_p = p^{-v_p(m)}. \tag{19}$$

由关系(15),(16),(17)得到:对 $m\neq0, n\neq0$,有

$$|m|_p = 1, \text{如果 } m\neq 0 \text{ 且 } p \text{ 不整除 } m; \tag{20}$$

$$|-m|_p = |m|_p; \tag{21}$$

$$|mn|_p = |m|_p \cdot |n|_p. \tag{22}$$

关系式(21),(22)显然当 $m=0$ 或 $n=0$ 时也为真. 最后,对所有整数 m,n,有

$$|m+n|_p \leqslant \max\{|m|_p, |n|_p\} \leqslant |m|_p + |n|_p. \tag{23}$$

按定义,此式当 $m=0$ 或 $n=0$ 或 $m+n=0$ 时显然成立;在别的情形,它是(18)的推论. [202]

最后,在所有整数的集合 **Z** 上可由公式

$$d_p(m,n) = |m-n|_p \tag{24}$$

定义一个距离. 距离公理的验证直接来自定义和关系式(21),(23). $d_p(m,n)$称为 m,n 之间的 p 进距离.

p 进距离具有与通常距离 $|m-n|$ 差异极大且看起来十分奇怪的性质. 如果非零整数 a 不被 p 整除,则对所有 $n\geqslant 0$ 有 $|a^n|_p=1$. 另一方面,当 n 趋于 $+\infty$ 时, $|p^n|_p = p^{-n}$ 收敛于 0. 这还表明,对于两个不同的素数 p,q,距离 d_p 与 d_q 是不等价的.

5. 傅里叶级数

A. 三角级数和傅里叶系数

我们假定已知区间$[a,b]$上连续函数级数

$$f_0(x) + f_1(x) + \cdots + f_n(x) + \cdots \tag{1}$$

的两个基本性质:

(I) 如果 $A_0 + A_1 + \cdots + A_n + \cdots$(对所有 n, $A_n \geqslant 0$)是收敛数

项级数，且对所有 $x\in[a,b]$ 有 $|f_n(x)|\leqslant A_n$，则级数(1)在 $[a,b]$ 上一致收敛，从而其和 $f(x)$ 在 $[a,b]$ 上连续.

(ii) 在同样假定下，有

$$\int_a^b f(x)\mathrm{d}x=\int_a^b f_0(x)\mathrm{d}x+\int_a^b f_1(x)\mathrm{d}x+\cdots+\int_a^b f_n(x)\mathrm{d}x+\cdots, \tag{2}$$

其中右边级数绝对收敛，且

$$\int_a^b |f(x)|\mathrm{d}x\leqslant(b-a)(A_0+A_1+\cdots+A_n+\cdots).$$

形如

$$a_0+(a_1\cos2\pi x+b_1\sin2\pi x)+\cdots+(a_n\cos2\pi nx+b_n\sin2\pi nx)+\cdots \tag{3}$$

的级数称为周期为 1 的*三角级数*. 如果此级数在点 $x\in\mathbf{R}$ 处收敛，则它对所有点 $x+m$ 收敛(m 是整数)，且具有同样的和. 如果它对所有点 $x\in[0,1]$ 收敛，则它在 $\mathbf{R}$ 上收敛，且其和 $f(x)$ 是周期为 1 的*周期函数*，即 $f(x+1)=f(x)$.

我们特别来考虑级数

$$|a_0|+(|a_1|+|b_1|)+(|a_2|+|b_2|)+\cdots+(|a_n|+|b_n|)+\cdots \tag{4}$$

收敛的情形. 由于 $|\cos2\pi nx|\leqslant1$，$|\sin2\pi nx|\leqslant1$，所以级数(3)绝对并一致收敛，且其和 $f(x)$ 在 $\mathbf{R}$ 上连续并具有周期 1. 此外，[203] 值得注意的是，可以借助 $f(x)$ *表示系数* a_k 和 b_k. 这是下面的称为正交性公式的推论：

对所有整数 m,n，

$$\int_0^1\cos2\pi mx\cdot\sin2\pi nx\,\mathrm{d}x=0; \tag{5}$$

对整数 $m,n,m\neq n$，

$$\int_0^1\cos2\pi mx\cos2\pi nx\,\mathrm{d}x=0,\ \int_0^1\sin2\pi mx\sin2\pi nx\,\mathrm{d}x=0; \tag{6}$$

对 $m\neq0$，

$$\int_0^1 \cos^2 2\pi mx\,\mathrm{d}x = \frac{1}{2},\ \int_0^1 \sin^2 2\pi mx\,\mathrm{d}x = \frac{1}{2}. \tag{7}$$

这些公式可直接从经典三角公式

$$\cos a \sin b = \frac{1}{2}(\sin(a+b) + \sin(b-a)),$$

$$\cos a \cos b = \frac{1}{2}(\cos(a+b) + \cos(a-b)),$$

$$\sin a \sin b = -\frac{1}{2}(\cos(a+b) - \cos(a-b))$$

以及下列事实推出：对非零整数 p，

$$\int_0^1 \cos 2\pi px\,\mathrm{d}x = 0,$$

$$\int_0^1 \sin 2\pi px\,\mathrm{d}x = 0.$$

于是在级数(4)收敛性的假定下，就有借助 $f(x)$ 给出(3)的系数的傅里叶公式：

$$\begin{aligned} a_0 &= \int_0^1 f(x)\,\mathrm{d}x, \\ a_n &= 2\int_0^1 f(x)\cos 2\pi nx\,\mathrm{d}x,\ n \geqslant 1, \\ b_n &= 2\int_0^1 f(x)\sin 2\pi nx\,\mathrm{d}x,\ n \geqslant 1. \end{aligned} \tag{8}$$

事实上，作为性质(ii)的推论，为计算积分

$$\int_0^1 f(x)\cos 2\pi nx\,\mathrm{d}x,$$

以 $\cos 2\pi nx$ 乘级数(3)，它能逐项积分，其结果可从公式(5)，(6)，(7)得到. 积分

$$\int_0^1 f(x)\sin 2\pi nx\,\mathrm{d}x$$

能以同样方法计算.

对[0,1]上每个连续函数 $f(x)$（但不假定它是级数(3)的和），[204] 由公式(8)定义的系数 $a_0, a_n, b_n (n \geqslant 1)$ 称为 f 的傅里叶系数.

当 f 具有连续导数 f' 时，也可确定 f' 的傅里叶系数：

$$
\begin{aligned}
a_0' &= \int_0^1 f'(x)\mathrm{d}x, \\
a_n' &= 2\int_0^1 f'(x)\cos 2\pi nx\,\mathrm{d}x, n \geqslant 1, \\
b_n' &= 2\int_0^1 f'(x)\sin 2\pi nx\,\mathrm{d}x, n \geqslant 1.
\end{aligned} \tag{9}
$$

而且这些系数与 f 的傅里叶系数之间有下列关系：

$$
\begin{aligned}
a_0' &= f(1) - f(0), \\
a_n &= -\frac{1}{2\pi n}b_n', n \geqslant 1, \\
b_n &= \frac{1}{2\pi n}(a_n' - 2a_0'), n \geqslant 1.
\end{aligned} \tag{10}
$$

这些关系可简单地通过分部积分来证明，例如，

$$
\begin{aligned}
a_n &= 2\int_0^1 f(x)\cos 2\pi nx\,\mathrm{d}x \\
&= \left[\frac{1}{\pi n}f(x)\sin 2\pi nx\right]_0^1 - \frac{1}{\pi n}\int_0^1 f'(x)\sin 2\pi nx\,\mathrm{d}x.
\end{aligned}
$$

B. 傅里叶级数的收敛性

对[0,1]上每个连续函数 f，一当我们构作 f 的傅里叶系数(8)后，就能考虑相应的三角级数(3)．我们称它为 f 的傅里叶级数．立即提出两个问题：此级数是否收敛？如果它在点 x 处收敛，其和是否为 $f(x)$？

当仅仅假定 f 为连续时，这是十分困难的问题．人们一直等到 1873 年，才发现连续函数其傅里叶级数在某些点不收敛的例子；直到 1966 年，才能证明这些点仅构成某种意义上“可忽略的”集合，而 f 的傅里叶级数在其余点处收敛并具有和 $f(x)$．

我们将限于级数(4)(其项为 f 的傅里叶系数的绝对值)收敛的情形．此时如我们在 A)中所见，傅里叶级数(3)对每个

$x\in\mathbf{R}$收敛，且其和为一连续周期函数 $g(x)$. 还剩下证明 $g=f$.

我们应用从几何级数直接引出的辅助级数

$$\frac{1}{1-z}=1+z+z^2+\cdots+z^n+\cdots, \tag{11}$$

上式对所有满足 $r=|z|<1$ 的复数 z 成立. 如果令

$$z=r(\cos t+i\sin t),$$

则有

$$z^n=r^n(\cos nt+i\sin nt)$$

(这是科茨和棣莫弗于 18 世纪初证明的公式). 取(11)两边的实部，得关系式

$$\frac{1-r^2}{1+r^2-2r\cos t}=1+2(r\cos t+r^2\cos 2t+\cdots+r^n\cos nt+\cdots), \tag{12}$$

[205] 右边的级数(对固定的 $r<1$)在 $\mathbf{R}$ 上一致收敛. 对 $0<r<1$，令

$$\begin{aligned}\Phi_r(x)&=\frac{1-r^2}{1+r^2-2r\cos 2\pi x}\\&=1+2\sum_{n=1}^{\infty}r^n\cos 2\pi nx.\end{aligned} \tag{13}$$

它是在 $\mathbf{R}$ 上连续、周期为 1 的周期函数，并具有下列性质：

(i) 对所有 $x\in\mathbf{R}$, $\Phi_r(x)>0$. 事实上，作为上面(11)式后所作演算的推论，(13)中的分母不是别的，而是

$$|1-r(\cos 2\pi x+i\sin 2\pi x)|^2.$$

(ii) 对每个满足 $0<\delta<1/2$ 的数 δ 和满足 $\delta\leqslant x\leqslant 1-\delta$ 的所有 x，有 $\cos 2\pi x\leqslant\cos 2\pi\delta$，因而

$$\begin{aligned}1+r^2-2r\cos 2\pi x&\geqslant 1+r^2-2r\cos 2\pi\delta\\&=(1-r)^2+4r\sin^2\pi\delta\\&\geqslant 4r\sin^2\pi\delta.\end{aligned}$$

于是当 $\delta\leqslant x\leqslant 1-\delta$ 时，有

$$\Phi_r(x)\leqslant\frac{1-r^2}{4r\sin^2\pi\delta}\leqslant\frac{1-r}{2r\sin^2\pi\delta}. \tag{14}$$

(iii) $$\int_0^1 \Phi_r(x)\mathrm{d}x = 1.$$

事实上,(13)右边的级数可逐项积分,而除第一项外所有积分都等于零.

建立这些结果后,让我们构作下面的函数:

$$F_r(x) = \int_0^1 \Phi_r(x-t)f(t)\mathrm{d}t. \tag{15}$$

对所有 $x \in \mathbf{R}$ 和所有 $t \in [0,1]$,有

$$\Phi_r(x-t)f(t) = f(t) + 2\sum_{n=1}^{\infty} r^n f(t)\cos 2\pi n(x-t). \tag{16}$$

由于 $|f(t)| \leqslant M$①,我们有

$$|r^n f(t)\cos 2\pi n(x-t)| \leqslant Mr^n,$$

因而(16)右边的级数一致收敛,从而能逐项积分.我们得到

$$F_r(x) = \int_0^1 f(t)\mathrm{d}t + 2\sum_{n=1}^{\infty} r^n \int_0^1 f(t)\cos 2\pi n(x-t)\mathrm{d}t. \tag{17}$$

由于

$$\cos 2\pi n(x-t) = \cos 2\pi nx \cdot \cos 2\pi nt + \sin 2\pi nx \cdot \sin 2\pi nt,$$

由(8)得

$$2\int_0^1 f(t)\cos 2\pi n(x-t)\mathrm{d}t = a_n \cos 2\pi nx + b_n \sin 2\pi nx,$$

从而最后有

$$F_r(x) = a_0 + \sum_{n=1}^{\infty} r^n (a_n \cos 2\pi nx + b_n \sin 2\pi nx). \tag{18}$$ [206]

这样,如果证明下列 1),2),关系

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos 2\pi nx + b_n \sin 2\pi nx) \tag{19}$$

就得证:

1) 当 r 趋于 1 时,(18)右边的级数趋于 f 的傅里叶级数.

2) 当 r 趋于 1 时,对所有 $x \in [0,1]$,$F_r(x)$ 趋于 $f(x)$.

---

① [0,1]上所有连续函数都是有界的.

1)的证明.(18)与(19)右边级数之差为

$$\sum_{n=1}^{\infty}(1-r^n)(a_n\cos 2\pi nx+b_n\sin 2\pi nx). \tag{20}$$

给定一个数 $\varepsilon>0$,存在自然数 N,使得

$$\sum_{n=N+1}^{\infty}(|a_n|+|b_n|)<\varepsilon,$$

因之

$$\left|\sum_{n=N+1}^{\infty}(1-r^n)(a_n\cos 2\pi nx+b_n\sin 2\pi nx)\right|\leqslant 2\varepsilon. \tag{21}$$

另一方面,对 $1\leqslant k\leqslant N$,有

$$1-r^k=(1-r)(1+r+r^2+\cdots+r^{k-1})\leqslant N(1-r).$$

如果令

$$A=\sum_{n=1}^{\infty}(|a_n|+|b_n|),$$

则有

$$\left|\sum_{n=1}^{N}(1-r^n)(a_n\cos 2\pi nx+b_n\sin 2\pi nx)\right|\leqslant AN(1-r). \tag{22}$$

于是从(21),(22)得知,对于级数(20),有

$$\left|\sum_{n=1}^{\infty}(1-r^n)(a_n\cos 2\pi nx+b_n\sin 2\pi nx)\right|\leqslant AN(1-r)+2\varepsilon. \tag{23}$$

这样,当 r 充分接近 1 以致 $1-r\leqslant\varepsilon/AN$ 时,(23)的右边的上界为 3ε,这就完成了证明.

2)的证明.固定一个数 $x\in[0,1]$,并假定给定一个 $\varepsilon>0$.由于 f 连续,存在数 δ,使得 $0<\delta<1/2$ 且对 $x-\delta\leqslant t\leqslant x+\delta$ 有

[207]

$$|f(x)-f(t)|\leqslant\varepsilon. \tag{24}$$

(如果 $x-\delta<0$ 或 $x+\delta>1$,则必须分别在上述条件中以 0 代替 $x-\delta$ 或以 1 代替 $x+\delta$.)注意,作为 $\Phi_r(x)$ 的性质(iii)的推论,我们能写

$$f(x)=\int_0^1\Phi_r(x-t)f(x)\,\mathrm{d}t.$$

为证实这一点，应用下述事实：对每个在 **R** 上连续且周期为 1 的周期函数 h，

$$\int_0^1 h(t)\mathrm{d}t = \int_a^{a+1} h(t)\mathrm{d}t \tag{25}$$

对每个 $a \in \mathbf{R}$ 成立. 如果 $a = n$ 是整数，只须注意由周期性，变量替换 $t = n + u$ 给出

$$\int_n^{n+1} h(t)\mathrm{d}t = \int_0^1 h(n+u)\mathrm{d}u = \int_0^1 h(u)\mathrm{d}u.$$

如果对某个整数 n，$n < a < n + 1$，则

$$\int_a^{a+1} h(t)\mathrm{d}t = \int_a^{n+1} h(t)\mathrm{d}t + \int_{n+1}^{a+1} h(t)\mathrm{d}t,$$

通过变量替换 $t = y + 1$ 得

$$\int_{n+1}^{a+1} h(t)\mathrm{d}t = \int_n^a h(y+1)\mathrm{d}y = \int_n^a h(t)\mathrm{d}t,$$

这就给出(25).

由(25)得

$$f(x) - F_r(x) = \int_0^1 \Phi_r(x-t)(f(x) - f(t))\mathrm{d}t.$$

为简单起见，写 $g(t,x) = \Phi_r(x,t)(f(x) - f(t))$. 然后有

$$\int_0^1 g(t,x)\mathrm{d}t = \int_0^{x-\delta} g(t,x)\mathrm{d}t + \int_{x-\delta}^{x+\delta} g(t,x)\mathrm{d}t + \int_{x+\delta}^1 g(t,x)\mathrm{d}t. \tag{26}$$

(如果 $x - \delta < 0$ 或 $x + \delta > 1$，则分别取消第一或第三项，而在第二项中分别用 0 代替 $x - \delta$ 或用 1 代替 $x + \delta$.)

我们分别考虑(26)中三项积分的上界. 已知 $\Phi_r(x-t) > 0$，故由(24)得到

$$\left|\int_{x-\delta}^{x+\delta} \Phi_r(x-t)(f(x) - f(t))\mathrm{d}t\right| \leqslant \epsilon \int_{x-\delta}^{x+\delta} \Phi_r(x-t)\mathrm{d}t \leqslant \epsilon,$$

因为

$$\int_{x-\delta}^{x+\delta} \Phi_r(x-t)\mathrm{d}t \leqslant \int_0^1 \Phi_r(x-t)\mathrm{d}t = 1.$$

另一方面,作为(14)的推论,有

$$\left|\int_0^{x-\delta}\Phi_r(x-t)(f(x)-f(t))\mathrm{d}t\right|\leqslant 2M\frac{1-r}{2r\sin^2\pi\delta},$$

[208] 对于展布在 $x+\delta$ 到 1 上的积分,也有同样的上界.

最后,如果 $r>1/2$,则有不等式

$$|f(x)-F_r(x)|\leqslant\epsilon+2M\frac{1-r}{r\sin^2\pi\delta}\leqslant\epsilon+4M\frac{1-r}{\sin^2\pi\delta}. \quad (27)$$

如果 r 接近 1 致使 $r>1/2$ 与

$$1-r\leqslant\frac{\epsilon\sin^2\pi\delta}{4M},$$

则 $|f(x)-F_r(x)|\leqslant 2\epsilon$,从而(19)得证.

C. 伯努利多项式的傅里叶级数

我们记得(第四章附录 4),对每个自然数 n,已定义 n 次伯努利多项式 $\varphi_n(x)$. 我们提一下它的一些主要性质:

对 $k\geqslant 1$, $\quad\varphi_{2k+1}(0)=\varphi_{2k+1}(1)=0,$

$$\varphi_{2k}(0)=\varphi_{2k}(1)=(-1)^{k+1}B_k; \quad (28)$$

对 $n\geqslant 2$, $\quad\varphi'_n(x)=n\varphi_{n-1}(x). \quad (29)$

最后,我们有

$$\varphi_1(x)=x-\frac{1}{2},$$

所以 $\varphi'_1(x)=1$.

这些公式以及公式(10)直接确定 φ_n 的傅里叶系数. 对于 φ_1,有

$$a_n=0,n\geqslant 0;b_n=-\frac{1}{\pi n}. \quad (30)$$

对于所有 φ_n,我们有

$$a_0=\int_0^1\varphi_n(x)\mathrm{d}x=0,$$

因为它可写为

$$\frac{1}{n+1}\int_0^1 \varphi'_{n+1}(x)\mathrm{d}x=\frac{1}{n+1}(\varphi_{n+1}(1)-\varphi_{n+1}(0))=0,$$

最后的等式是由于(28)式.

另外的傅里叶系数由(30)以及用公式(10)关于 n 归纳地推出.对 $\varphi_{2k}(x)$,我们有

$$a_n=(-1)^{k+1}\frac{2(2k)!}{(2\pi n)^{2k}}, b_n=0;$$

而对 $\varphi_{2k+1}(x)$,则有

$$a_n=0, b_n=(-1)^{k+1}\frac{2(2k+1)!}{(2\pi n)^{2k+1}}.$$

由于对于从 $m\geqslant 2$ 起的所有多项式 $\varphi_m(x)$,级数

$$\sum_{n=1}^{\infty}(|a_n|+|b_n|)$$

收敛,我们就容易地得到了第四章附录 4 中的方程(14)和(15).

D. 康托尔问题

可能发生这样的情形:三角级数(3)对所有 $x\in\mathbf{R}$ 收敛(但不绝对收敛),但其和不一定连续,因而积分(8)可能没有意义. [209] 例如,级数

$$\sum_{n=2}^{\infty}\frac{1}{\sqrt{\log n}}\sin 2\pi nx$$

就是这种情形.因此三角级数概念比傅里叶级数概念更一般.这是由黎曼首次考虑的,他证明了下述定理:如果级数(3)对所有 x 收敛且其和为零,则所有系数 a_n, b_n 都为零.康托尔试图知道,如果假定级数(3)除一个集合 E 的点之外收敛且其和为零,上述结论是否仍为真.正是这个问题把他引向 $\mathbf{R}$ 的任意子集的研究,特别从这些子集的序结构或拓扑的观点来研究.但是他很快放弃了最初提出的问题,而它至今仍未完全解决. [210]

第 六 章

关于"数学基础"的问题和假问题

我想我可以说，前面三章所作的阐述贯彻了我在引言中概述的纲要：表明当代数学对象的"抽象"性质如何来自以解决经典问题为目的而于1800至1930年间发明的方法，并且引出这样的认识——它由于我不能过多进入技巧细节而受到限制——即如今这些方法以及它们所应用到的对象的效用比以往任何时候都要巨大．

我们从数学的"假设－演绎"观念开始，柏拉图已描述过这一观念，而欧几里得则在他的《几何原本》中把它付诸实践．然后，可以这么说，我们跃过许多个世纪，看到帕施和希尔伯特如何对几何学给出这个观念的一种完全精确的形式，它几乎立刻被他们的同辈和直接继任者推广到数学的所有分支，以至可以说它已成为当今数学家的常规操作手段．

在这样做的时候，我们略去了1800至1930年这个阶段的整个另一方面并对它保持缄默，这就是：这是一个——正如我们所强调的——在创造新理论方面异常富有成果的时代，但也是关于数学本性的怀疑和争执达到高峰的时代．现在除了历史学家和哲学家外，由这些怀疑和争执引起的辩论已很少有人感到兴趣．然而，这些辩论以17和18世纪关于"无穷小"的争论同样的方式，在数学语言不断增长的精确化方面留下了正面的效果．特别是，它为数理逻辑在20世纪中的巨大进步开辟了道路．

[211] 如我们在第五章开头所说，正是数学家关于数学对象本性的认识在1800年以后发生了变化．许多数学家已不像他们的前

辈那样满足于运用17世纪中锻造的工具所能取得的巨人般的前进步伐，并认为它足以确定这些方法所依据的各种原理的有效性，而是提出其“真实性”已被宣称为确实无疑的这些原理的真正意义何在的问题．如果我想用一句话来概括这一阶段各种观念卷入的方式，那么我要说，它的本质在于首先在几何学然后在数学的其余部分逐步抛弃“显见为真”的观念．

1．非欧几何学

A．平 行 公 设

我们说过(第三章，§4)，17和18世纪的数学家关于几何公设的态度与柏拉图和欧几里得的态度极不相同，在后者仅仅谈到“假设”或“规定”的地方，前者要在其中看出“真理”，而不去解释这种特征如何能在应用于——如同罗巴切夫斯基所说——“自然界中不存在的”概念的命题中看出；他们正确地指责“几何学一开始所用的语汇”就是含糊不清的．用康德的说法，这些公设是“先验综合判断”，它们可使哲学家满意，但无助于数学家精确表达他们的想法．

这样，在接受由几何公设描绘的几何对象的性质与它们在自然界中“形象”的性态之间完美和谐方面有着普遍的共识．例如，对于一条直线能无限延长这个断言，除了显然性外，看来没有什么人再能找出别的东西；或许这是由于它对应于人们“近似地”并在非常有限的范围内以实验方式所能做的事情①．

① 这或许也由于，虽然这些公设本身没有令人信服的“原始”形象，但因其逻辑推论提供了与经验十分相符的自然科学模型，因此其“真理性”毫无困难地被接受．关于“质点”动力学的牛顿定律也有类似的情形，虽然它们本身的存在性多多少少难以证明，但却能从中导出大量完全满意地“解释”现象的结论．

然而,欧几里得有一个"规定",即使在古典时代也已成为讨[212]论的主题,这就是著名的"平行公设",它最后被称为"欧几里得公设".在平面上给定一个不在直线 Δ 上的点 A,容易作出一条过点 A 且平行于 Δ 的直线 Δ':只须(如图 45)作垂直于 Δ 的直线 AB,再在点 A 处作垂直于 AB 的直线 Δ'.直线 Δ' 不可能与 Δ 相交,因为否则两条垂直于 AB 的不同直线就会通过 Δ 与 Δ' 的交点,而这同先前的定理(欧几里得,《几何原本》,第Ⅰ卷,命题 27)相矛盾.

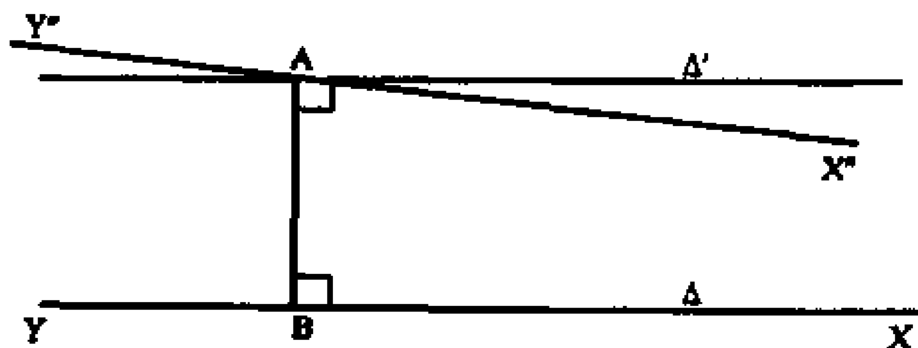


图 45

欧几里得公设所断言的是:Δ' 是通过点 A 且平行于 Δ 的唯一的直线;至少如今它是这样陈述的.欧几里得以比较肯定的形式表述这一性质:如果在由直线 D 的一侧的点构成的半平面中,两条半直线 AX_1,BX_2 分别与 D 的线段 AB 构成角 α,β,且 α,β 之和小于两直角(图 46),则这两条半直线具有交点(《几何原本》,第Ⅰ卷,公设 5).

图 45 中过点 A 平行于 Δ 的直线的唯一性可由上述公设推出:过点 A 且使 $\angle BAX''$ 小于直角的直线 $Y''X''$ 与 BX 相交,因为 $\angle BAX''$ 与 $\angle ABX$ 之和小于两直角.如果 $\angle BAX''$ 大于直角,则[213]可对半直线 AY'' 与 BY 作相同的论证.

尽管欧几里得这里涉及的仍是等同于直线无限延长的肯定论断,但因这是接受位于任意远距离的点的存在性这一事实,所

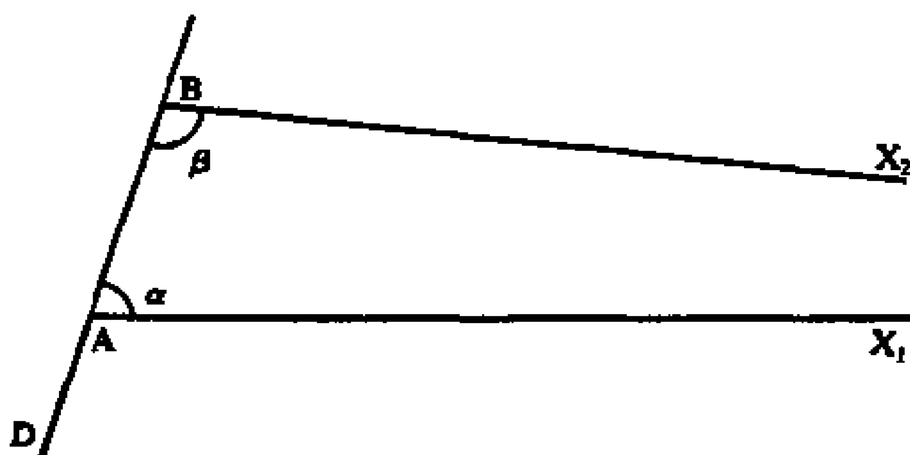


图 46

以他在建立他的主要定理时不能没有它，例如关于矩形的存在性——每个人都倾向于把它看作“真理”——就是如此．事实上（图 47），为作出以给定的3个点 A，B，C（$\angle ABC$ 为直角）为顶点的矩形，可在点 C 处垂直于 BC 的直线上取长度等于 AB 的线段 CD，然后必须证明 $\angle BAD$ 和 $\angle CDA$ 是直角．由于 $\angle ABC$ 与 $\angle BCD$ 相等，且由于 $AB = CD$，所以三角形 ABC 与三角形 BCD 全等，因而 $AC = BD$．三角形 BAD 与 CDA 的对应边相等，故 $\angle BAD$ 与 $\angle CDA$ 也相等；但为什么这两个角是直角？如果接受欧几里得公设，则可论证如下：如果 $\angle BAD$ 是锐角，则半直线 AX' 与 BX 相交，又由于 $\angle CDA$ 也是锐角，所以半直线 DY' 与 CY 也相交；这样直线 XY 与 $X'Y'$ 就会有两个交点，这是谬误的．

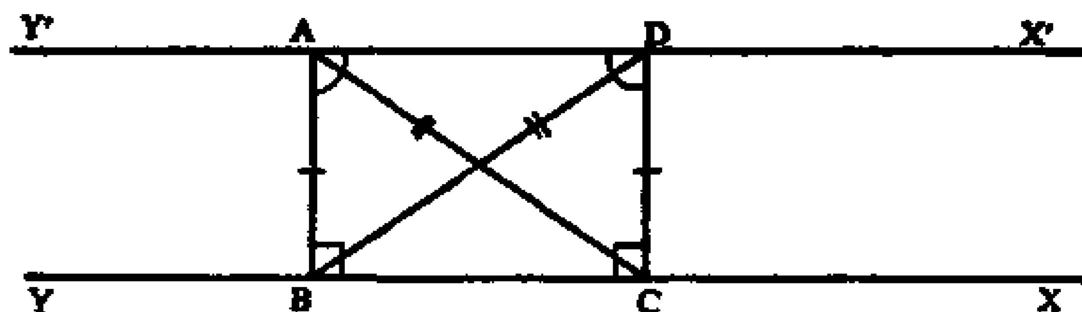


图 47

如果$\angle BAD$是钝角,则$\angle BAY'$和$\angle CDX'$是锐角,于是同样的论证又导致矛盾.

从古典时代就已出现的反对上述证明的意见,并不涉及它的正确性,而是涉及在理解欧几里得为何必须在这里引入一个新公设方面出现的错误,即认为应当可能仅用别的公理来证明该公设.在一千多年的历史进程中,从普洛克鲁斯到勒让德,许多数学家试图作出这种证明,但都以失败告终.在18世纪,萨开里于1733年,朗伯于1770年前后开创了新的策略:用反证法,[214]证明(不用欧几里得公设)$\angle BAD$和$\angle CDA$为锐角的假定[①]导致谬论.他们由这一假定开发出一长列与经典定理大相径庭的结论,例如,一个三角形的三内角之和再也不会等于两直角,而会依赖于所给三角形的面积.所有这些奇怪的定理全都与任何不依赖于平行公设的欧几里得定理不相矛盾.

在1795到1830年中,这些不断重复的失败终于导致到这样的结果:有几位数学家相信,他们所寻求的矛盾永远不会找到,在数学中留有几种不同几何学的余地.但这种信念与当时知识界普遍接受的看法完全相反,这种看法授予欧几里得几何学与我们的空间概念不可分割的必然品性.高斯从年青时就对平行线问题感兴趣,到大约1816年时,他已形成这样的想法,即如果我们接受与欧几里得公设相反的陈述,把它作为(柏拉图意义下的)"假设",那么就有可能建立一种他称之为"非欧几里得的"几何学.他判断尽管这种几何学具有奇怪的性质,但却"与自身完全相容".他把自己的思索告诉了一些朋友,但拒绝发表,对催促他发表的人他回答道:"我十分担心,如果我和盘托出自己的观点,那些愚人们就要穷嚷嚷."

稍后,两位数学家,匈牙利的J·波尔约和俄罗斯的N·罗巴

① 可以证明这两个角为钝角的假定与阿基米德公理(见第三章附录2)相矛盾.

切夫斯基各自独立地得到了高斯的结论，两人都大约于1830年发表了按欧几里得模型展开的关于"非欧几里得"几何学的论著．不过，在大约1860年之前，大多数数学家并不知道这些著作，而知道的数学家也只把它们判定为奇幻的，没有任何严肃的内蕴，不是数学的一个真实部分．

B．曲面上的几何学

1860年后，在微分几何的影响下，情况发生了变化．无庸置疑，高斯（受到他本人关于大地测量和天文学的许多工作的激励）是第一位蕴育出这样想法的数学家：在充分正则[①]的任一曲面Σ上，一旦知道如何定义画于Σ上的一条充分正则的曲线的弧段的长度，就有可能在Σ上构造一种"几何学"．事实上，人们能在这里定义相当于欧几里得平面上的直线的对象，它称为Σ上的测地线．这样的曲线Γ为下述性质所刻画：对于Γ上两个互相充分接近的点p,q，Γ的位于p与q之间的弧段在Σ上任何以p,q为端点的曲线弧中是最短的．例如，球面上的测地线是大圆，它是球面与通过球心的平面的交线．在直圆柱面上，测地线是： [215]

1）（铅直）母线；

2）垂直于所给圆柱面的轴的平面上由该圆柱面截出的圆；

3）可在该圆柱面上画出的螺旋线．

知道了画于Σ上的曲线的长度，首先就使我们能定义相交于一个点的两条测地线之间的角特别是定义直角（附录1）．不过，在试图按欧几里得模型在Σ上展开一种"几何学"时，我们几乎总要把自己限制在一个小区域中，因为只有在充分小的集合Δ中，才存在连接Δ的任意两点p,q且完全包含于Δ内的一条且仅存在一条测地弧．然后此测地线就能有效地定出Σ上

① 其精确定义见附录1．

连接 p 与 q 的曲线弧段长度的最小值①.这样,在 Δ 中,欧几里得的第一条"规定"(见第三章,§4)得到满足;第二条关于测地线段的"延长"也一样,只要我们不打算作"任意的"延长②.包含于 Δ 中的充分正则的图形的"面积"也能加以定义.

可是,一般地说,Δ 上的"几何学"必须到此为止.这里不再有"全等三角形的各种情形"③;例如,Δ 中的两三角形,可能具有相同长度的对应边,但对应的角却不相等.这是因为,一般地说,在 Δ 中不能有位移,即保持测地线、长度和角度不变的变换.

然而在某些(不同于平面的)曲面上,确实存在这样的变换.这些曲面具有传递性,这意味着 Δ 中测地线弧 pq 能通过这样的变换之一映为任一具有同样长度的别的测地线弧 $p'q'$.例如球面就有这种情形,此处"位移"是绕过球心的一个轴的旋转;因[216]而在这里对于边长小于半个大圆的三角形,"全等的各种情形"仍然成立.对于圆柱面这一点同样成立,只要取 Δ 为,例如,圆柱面上这样的部分,它投影到底面上是基圆的一个弧但非整个基圆.

这里要注意球面与圆柱面之间的一个基本差别:在圆柱面上,三角形 ABC 三内角之和 $A+B+C$ 仍等于两直角;而在球面上,不管三角形取得多么小,其内角之和恒大于两直角,精确地说,我们有(角以弧度度量)

$$A+B+C-\pi=\frac{\text{三角形 } ABC \text{ 的面积}}{R^2},$$

其中 R 是所给球面的半径.

---

① 在圆柱面上,总有无穷多条测地线通过两个不同的点 p,q,除非这两个点位于所给圆柱面与一垂直于该圆柱面之轴的平面相截得到的圆上.

② 在大多数欧几里得命题中,并不需要把一条直线延长超过所研究图形中各个线段之长的长度.

③ 所说"三角形"总以测地弧为其边.

所有这些实质上从古典时代起都早已知道[1]，但看来它对平行线问题毫无影响，无疑这是因为球面和圆柱面都不具备连接任意两点的测地弧为唯一的性质，也是因为没有小区域 Δ 的“位移”群[2].

C. 非欧几何模型

看来古希腊人并未试图计算除圆之外任何平面曲线的长度．直到 17 世纪，才有可能通过微积分提供可以计算充分正则的平面曲线长度的公式；如同圆周长度那样，曲线长度定义为“内接”于所给曲线的折线之长当折线各段最大长度趋于零时的极限(图 48).



图 48

欧拉把计算平面曲线长度的公式推广到嵌在通常 3 维空间 $\mathbf{R}^3$ 中的“挠”曲线(附录 1)；当他和他的生于黎曼之前的继任者谈到画于一张曲面 Σ 上的一条曲线的长度时，他们不管此曲线所在的曲面，面把它直接看成嵌在 $\mathbf{R}^3$ 中来计算其长度． [217]

然而，高斯的研究表明，如果曲面 Σ 可以作保持画于 Σ 上曲线长度不变的“变形”(例如，“变形”直圆柱面的底曲线而不改变其长度)，则必须把“变形后的”曲面上的“几何”看作与原曲面上的“几何”相同；换言之，所给曲面嵌入于 $\mathbf{R}^3$ 中这一点不再出现于计算之中．

直到 1860 年后，才有人想到破除这种定义长度的一贯程序，并想出许多其他定义方法．例如，让我们考虑半球面 Σ，它由含有球心的平面圆盘 Δ 所界定．Δ 的每个点是球面上唯一一

① 球面上的几何学和三角学对于计算天体位置是必不可少的．

② 当进行“位移”的“合成”时，通常会越出区域 Δ．

个点的正交投影,后者是前者的“原象”(图 49).

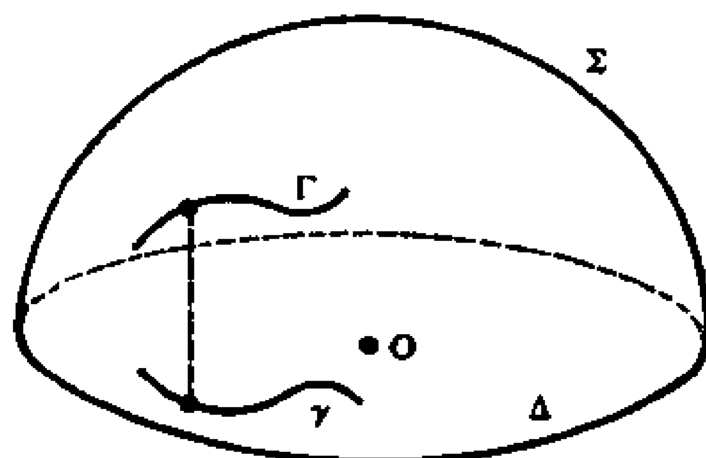


图 49

可以规定包含于 Δ 中的平面曲线 γ 的“长度”为其原象 Γ 的(欧拉意义下的)长度,面高斯的所有论证对于这种新的“长度”定义都成立.但这样得到的 Δ 中的几何与欧几里得几何有相当大的差别:测地线是 Σ 上的半大圆的正交投影,即长轴是 Δ 的一条直径的半椭圆,而“三角形”的“内角”之和恒大于两直角.

另一例子是在所谓“平坦环面”上定义长度.环面的“原始”形象可如下作出:取高为 h 的圆柱面,把它“弯曲”得使两端的两个圆粘在一起(图 50).可以毫无困难地在数学上定义此曲面

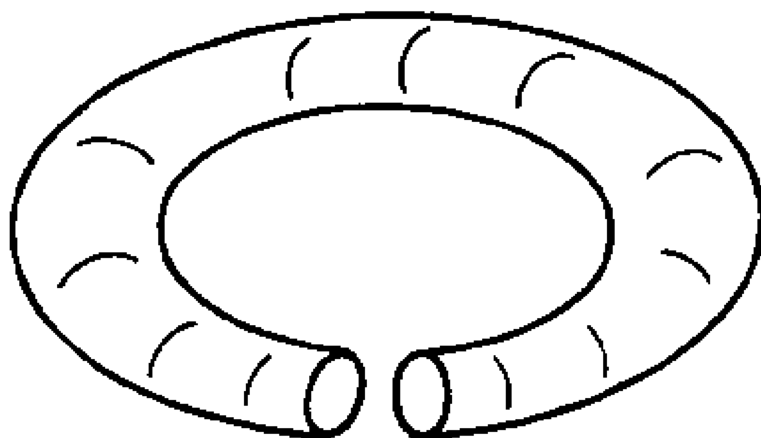


图 50

(附录 1),不过其上曲线 γ 的"长度"取为"弯曲"前圆柱面上对应曲线(欧拉意义下)的长度.在数学上这很容易定义(附录 1),但这导致与欧拉的定义完全不同的定义;例如,此环面上所有平行线具有相同的长度 h! 此时在该环面上的小区域中,其"几 [218] 何"与圆柱面上对应区域中的"几何"相同,特别是,三角形"内角"之和恒等于两直角.

或许这些例子看起来有点矫揉造作、稀奇古怪.它们仅仅是黎曼于 1854 年在其著名的"就职演讲"中引进的一般概念的非常特殊的情形;这一演讲直到黎曼谢世后的 1868 年才得以发表.对黎曼而言,曲面上的长度概念只遵从极少几个解析性质的条件(附录 1),而且他的想法前进得很远,因为他径直在他所称的"n 维流形"上定义长度,这里 n 是任一自然数.如同我们曾经说过的(见第五章, § 5, A, XVI),这些"黎曼空间"对我们的直观是如此陌生,却最终在当代物理学中起着主要作用.从我们现在关注的观点看,显然必须强调黎曼本来并无兴趣宣称他的公理的"显见真实性";他的就职演讲的标题,如同柏拉图那样,只提到"作为几何学基础的假设".

正是这种思想方法使贝尔特拉米[①]和克莱因于大约 1870 年能最终把可能仍然留存的对于"非欧几里得"几何合法性的疑虑一扫而光.要是这些几何的开创者把他们的论证再向前推进,他们就会碰到萨开里所寻求的矛盾,那么本来总是能反对他们的"信念"的.为证明相反的结论,贝尔特拉米和克莱因求助于一种理论在另一种理论中的模型这种观念.我们记得,最早这样的"模型"是人们通过下述途径来建立坐标方法:提供一本"词典",用它来把欧几里得平面几何中的任何对象或关系"翻译"为由实 [219] 数偶 (x, y) 构成的对象或关系,使得两种理论之一的任何定理对应于另一个中的相应定理(第三章, § 8).在现在所说的情形,

① 看来贝尔特拉米并不知道黎曼的"就职演讲".

贝尔特拉米和克莱因是在欧几里得几何中构造波尔约和罗巴切夫斯基非欧几里得平面几何的模型.庞加莱对他们的模型作了些许修改,其"词典"如下所述.这里左边写的是"非欧几里得"说法,而右边写的是相应的欧几里得译文:

平面　庞加莱半平面 H(第五章,§3,F)

直线　中心在 Ox 轴上的半圆或平行于 Oy 的半直线

角　角

圆　圆(包含于 H 中的)

包含于 H 中的曲线的(非欧)长度可以只用解析公式定义(附录1),但可以说,在 H 中点 (x,y) 的邻域内,小线段的非欧几里得长度约等于相应的欧几里得长度除以 y,这意味着在 Ox 轴的邻域内,非欧长度变为无穷大.同样,在点 (x,y) 的邻域内,小三角形的非欧面积约等于它的欧几里得面积除以 y^2.关于非欧长度的测地线恰是中心位于 Ox 轴上的半圆和平行于 Oy 的半直线;在半圆上,弧 A_1A_2 的非欧长度等于

[220]

$$\log\tan\frac{\varphi_2}{2}-\log\tan\frac{\varphi_1}{2},$$

面在平行于 Oy 的半直线上,线段 B_1B_2 的非欧长度是

$$\log y_2-\log y_1$$

(参见图 51).

然后就能检验欧几里得诸公理(更好地说,希尔伯特诸公理),当然平行公设除外.我们注意到,所有这些公理的"译文"都是欧几里得几何中的定理(或者,如果我们喜欢,通过使用坐标成为实数理论的定理).例如,欧几里得的第一个"规定"翻译为下述定理:对于 H 的不位于 Oy 的平行线上的两个不同的点 A_1, A_2,有且仅有一个中心位于 Ox 上的半圆通过所给两点.它可简单地证明如下:作 A_1A_2 的中垂线 Δ,由于 A_1A_2 不平行于 Oy,所以 Δ 不平行于 Ox,因此与 Ox 交于一单个点 C,此点即为我们要求的半圆的中心(参见图 51).半圆上线段 A_1A_2 的"无限"

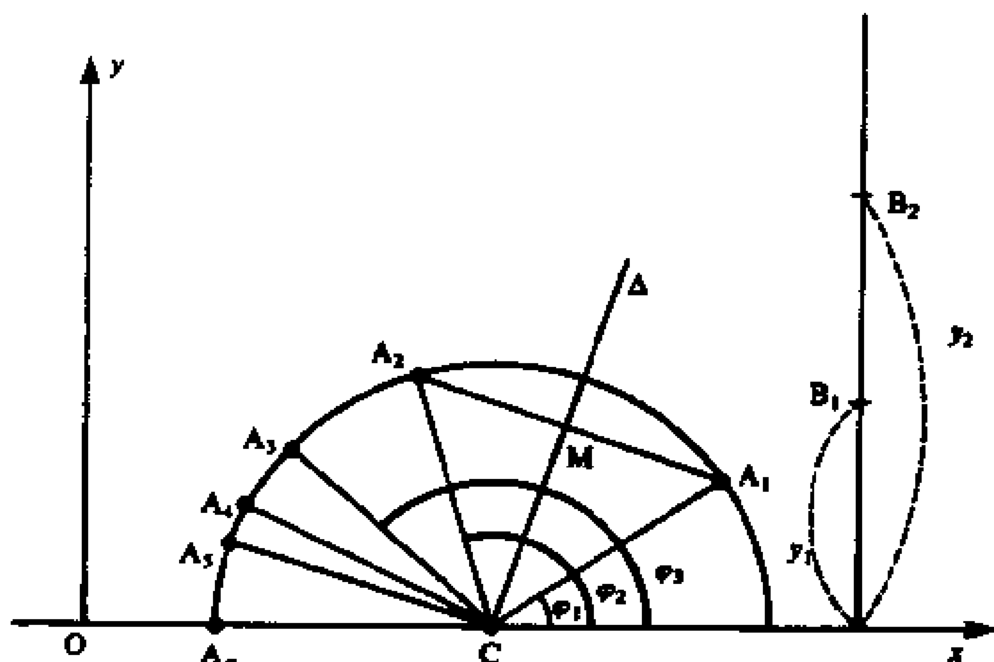


图 51

延长是可能的,甚至阿基米德公理也能在此半圆上加以验证,因为这一线段不断"加倍"得到端点 $A_2, A_3, \cdots, A_n, \cdots$,而

$$\log\tan\frac{\varphi_{n+1}}{2}-\log\tan\frac{\varphi_n}{2}=n\left(\log\tan\frac{\varphi_2}{2}-\log\tan\frac{\varphi_1}{2}\right).$$

因而当 n 无限递增时角 φ_n 趋于 π(但要注意,它永远不会达到 Ox 轴上的点 A_∞;该点不在 H 中).

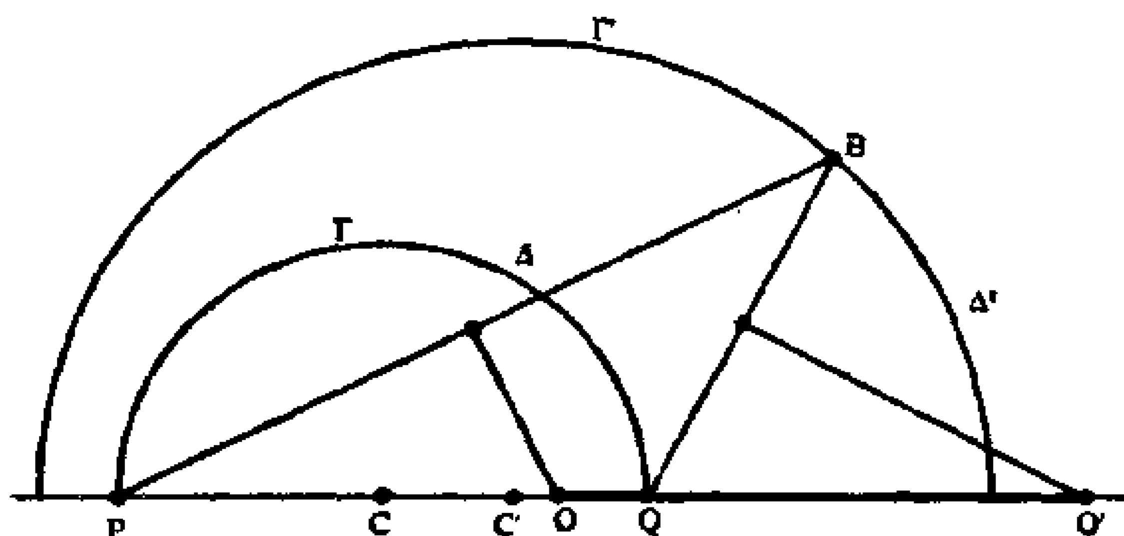
在继续演绎过程中具有决定意义的因素是非欧"位移"群的存在性,它就是我们在第五章 § 3,F 中谈到过的群 $GL^+(2,\mathbf{R})$;它具有使非欧长度和角度保持不变的性质(附录 1),并能仅在非欧距离 $A'B'$ 与非欧距离 AB 相等的条件下把 H 的两个点 A, B 映为另外两个点 A', B'. 由于存在这个群,就能验证,例如,"三角形全等的各种情形"在非欧意义下成立.

一旦所有"非欧几里得"公理翻译为欧几里得定理,我们就能安心地断定,从这些公理决不会导出矛盾,因为否则我们就会在"翻译"它们中在欧几里得几何里达到矛盾;而正如庞加莱所说,"无人怀疑通常几何没有矛盾".

另一方面,立即可以推出,对于非欧直线 Δ,在 H 中存在无穷多条过不在 Δ 上的点 B 且"平行"于 Δ 的直线;例如,如果 B [221] 在半圆 Δ 外侧,则只须考虑所有其圆心 C' 在 Ox 上并且过点 B 和 Ox 上不在 Δ 的直径 PQ 上的任一点的半圆(参见图 52).简单计算表明,对 H 中任一非欧三角形 ABC,其内角之和恒小于两直角,事实上我们有

$$\pi-(A+B+C)=ABC \text{ 的面积},$$

其中面积自然取非欧意义(附录 1).此外,容易看出和 $A+B+C$ 能要多小就有多小.也要注意,在波尔约-罗巴切夫斯基非欧几何中没有矩形,因为如果 Δ,Δ' 是两条非欧"平行"线,那么只有一条非欧直线同时垂直于 Δ 和 Δ'(附录 1).



(C'不能位于线段 OO' 上)

图 52

克莱因证明,也存在一种非欧几何,其中(如同球面)没有"平行线",且任一直线的"长度"为有界,然而总是只有一条测地线通过两个不同的点(这与球面上的情形相反).但是这样的几何定义于其上的曲面不可能嵌入到通常空间 $\mathbf{R}^3$ 中,它只能嵌

入到 $\mathbf{R}^4$ 中.

当相信"显见真实性"的人们不得不接受非欧几何逻辑上的合理性时,许多人仍然认为,经验能表明,只有欧几里得几何才能提供符合物理实在的空间模型. 庞加莱专门对此作了明确的论述,清楚指出这种想法是靠不住的:经验不能证明基于某种几何学的模型的"真实性",它只能证明该模型与某种物理理论相容;如果物理理论有了改变,那么必须改变几何模型以保持这种相容性([9]). [222]

2. 深入挖掘数的概念

A. 无 理 数

直到文艺复兴时代,在数学中,几何学对别的领域占有突出地位是无可竞争的①. 我们记得,欧几里得觉得有必要把自然数"表示"为线段,他并且用求解平面几何作图的形式来解释我们归结为二次方程的问题. 代数和分析从 17 世纪起的突飞猛进及其在几何学中的应用,以及欧几里得方式的"纯粹"几何学有点停滞的状态,导致观点的改变. "几何学的严格性"继续得到赞美,而从 19 世纪初叶起,对他们的前辈的不严格、不精密感到厌烦的分析学家,探索使无穷小分析具有他们在几何学中确认的一连串逻辑推理那种不可抗拒的特征,这就是所谓的"回到严格性",其中经历了作为分析基础的数列极限概念(见第三章, § 9)的建立. 然而,1817 年,波尔查诺——这一改革的杰出倡导者之一——认为,"纯粹(或一般)数学由算术、代数和分析组成"而几何只是一个"应用(或特殊)部分".

代数和分析的基础概念必定是数(即我们所用的实数). 事

① 这种突出性的一个证据是"几何学家"这个名称;直到 19 世纪末之前,它一直用来指所有数学家,甚至包括并不涉足几何学的数学家!

实上,柯西 1821 年的"分析教程"即从这个概念开始.他确实只限于说"数由量的绝对度量导出……并在前面置以 + 号或 - 号",并不加证明地列举了数的运算及其性质.这样,除了明确地陈述了对微积分的"严格"证明不可或缺的"区间套原理"(第三章,§6)外,他与笛卡儿和牛顿(他们的想法与欧几里得《几何原本》第V卷直接相联)停留在同一水平上.由于柯西并未给出"区间套原理"的任何证明,人们或许会想,他基于几何解释认为这是显然成立的;而当高斯不加证明地陈述与这个原理等价的有界序列最小上界存在性定理时,他必定也抱有同样的想法.但是[223]波尔查诺本人在 1817 年(在柯西之前)陈述了同一原理的一种形式,并认为求助于"几何定理的显然性"以证明数的性质,是"一种反对好方法的不可容忍的错误".他同样没有提供证明区间套序列公共数存在性的任何别的方法,而只限于肯定这一存在性"没有任何不可能之处"!

事实上,除几何学外,人们能从哪里找到无理数"存在性"的证明?古希腊人已能很好地定义无理数之间的关系(相等、不相等、加法,见第三章,附录 1),但他们认可无理数的存在性.从 1820 年起,认为应当置于经典数学基础之上的不是古希腊人的几何概念而是自然数概念这种想法开始出现.这就是 M·欧姆早于 1822 年就已在其雄心勃勃的"专论"①一开头所表达的想法,他打算把此书写成对整个数学相当于欧几里得《几何原本》的典籍.他在那里指责柯西没有在自然数基础上构造实数理论,然而他本人对此也无能为力.按戴德金所说,狄里克雷曾断言任何代数或分析定理"可陈述为关于自然数的定理".然而,直到 1860 年前后,才开始出现实数系的"构造".从我们的观点来看,它可以看成自然数理论中公理式实数理论(见第三章,附录 2)的模

① 此书书名为"Versuch eines vollkommen cosequenten Systems der Mathematik"(《数学的一个完备相容系的研究》).——译注

型，恰如同我们已经看到的在实数理论中构造的(欧几里得或非欧几里得)几何模型(见第三章，§8以及本章§1，C)．所有经典数学都以这种方式拥有自然数理论中的模型，这就是所谓数学的算术化．它在19世纪末取得如此成就，以至在学生进入大学前就要加以传授，这意味着他们不得不学习今后几乎毫无用处的论证．我们将限于在附录2中概略描述戴德金和梅雷-康托尔两种构造实数的基本原理．

B．怪　胎

尽管不能肯定其中有必然的因果关系，然而这些模型的构造必定使19世纪后半期的分析学家在当时他们所面对的可以称为“怪胎”的东西来临之前消除了疑虑，这里“怪胎”指的是这样的数学现象，它们对于人们的“直观”赋予的空间性质而言是完全不能预见和全然相反的．

[224]

年代最早的“怪胎”涉及曲线的切线．直到19世纪，“平面曲线”概念仍然不够精密．引进坐标系后，平面曲线看作由坐标之间的一个关系式 $F(x,y)=0$ 定义，或者，受到动点轨道作为“时间”t 的函数的启示，由一个“参数表示式”

$$x=f(t),y=g(t)$$

定义；但函数 F,f,g 必须具有没有精确规定的“正则性”，例如，由“初等函数”组成．无论如何，虽然没有明说，但看来总能通过铅笔在纸上的运动画出平面曲线，这意味着作这种运动时在每一点都有一个“方向”．数学上这被表达为曲线在每一点(可能除去有限个称为“角点”的点)存在切线．

在19世纪初，一旦连续函数概念得到精密表述，所讨论的问题就取下述形式：定义在 $\mathbf{R}$ 的一个区间中的连续函数 $y=f(x)$ 可能除有限个点外在该区间的每一点都具有导数吗？许多人不加讨论地承认是这样．柯西对这个问题没有作出断言，在绝大多数场合他只限于考虑假定其导数存在的函数．然而也有

一些人试图对此加以证明,其中包括安培,甚至也有还是高等师范学校学生的伽罗瓦.然而,在波尔查诺于1830年前后所写的一份手稿(它到一个世纪后才得以发表)中,给出了区间$[a,b]$上的一个连续函数,可以证明它在该区间的任一点处都没有导数①.直到1875年,才发表了在每个点都没有导数的连续函数的另一个例子,它是魏尔斯特拉斯于1872年开始的讲授中描述的.后来巴拿赫能够证明,在完全合理的意义下,这样的函数比可微函数"多得多"(第五章,§4,C).

这还仅仅是一个开头,而这类"反例"变得如此之多,以至可以把它们写成整整一本书.我们只限于在这类最为异乎寻常的例子中举出:在任何区间上都不等于常数的可微函数,它在每个区间中具有无穷多个极大、极小值;空间中的一张"曲面",它同[225]时是三个"房间"的通墙,这多少难以想象!但是,最奇怪的无疑是皮亚诺发明的"曲线":此"曲线"由两个连续函数

$$x=f(t),y=g(t),0\leqslant t\leqslant 1$$

定义,却填满了正方形$0\leqslant x\leqslant 1,0\leqslant y\leqslant 1$(换言之,对该正方形中任一点$(a,b)$,至少存在$t$的一个值$t_0$,使得$a=f(t_0),b=g(t_0)$).

这些例子表明了什么?在我看来,它们简单地表明,在我们对空间的信以为真的"直观理解"与定义几何学的数学对象的公理系统之间的任何联系都完全是表面的,这正好与17,18世纪数学家所想的相反.因此,在"真理"这个词通常所理解的意义下来谈论这些公理的"真理性",看来是错误的.

C. 算术的公理化

欧几里得并不认为需要对他的《几何原本》中关于算术的几

① 波尔查诺只说在任何区间(不管它多么小)中都有该函数没有导数的点.

卷给出公理式的表述，这多少有点奇怪；或许他相信，对于毫无异议地接受的基本概念，再引进“假设”是没有用处的？中世纪的坎帕纳斯(13 世纪)想填平这个缺口，但他的公理系统是不完备的．一直到大约 19 世纪中叶之前，算术书籍都只是追随欧几里得的写法．这个领域中第一位革新者是 H·格拉斯曼(1861 年)．他注意到，为在正整数(“自然数”)集 $\mathbf{N}^*$ 中定义加法和乘法，只须使用算子 $n \longmapsto n+1$；事实上，如果 $n+m$ 已定义，则我们可定义

$$n+(m+1)=(n+m)+1.$$

对于乘法，方法相同，取 $n\cdot 1=n$，而当 $n\cdot m$ 已定义，则令

$$n\cdot(m+1)=n\cdot m+n.$$

格拉斯曼由此出发证明了加法的结合性和交换性，乘法的结合性和交换性以及它关于加法的分配性．例如，为证明 $a+(b+c)=(a+b)+c$，他注意到由加法的定义，此命题对于 $c=1$ 为真；而如果 $c=x+1$ 而且如果我们假定已证明 $a+(b+x)=(a+b)+x$，则可有

$$\begin{aligned} a+(b+c) &= a+(b+(x+1))=a+((b+x)+1) \\ &= (a+(b+x))+1=((a+b)+x)+1 \\ &= (a+b)+(x+1)=(a+b)+c. \end{aligned}$$

我们注意到，正如帕斯卡时代以来的数学家所做的那样，格拉斯曼在这里隐含地使用了所谓*递推原理*(或 19 世纪中所称的“完全归纳原理”)，而没有陈述其一般形式，即：如果我们想对所有 [226] 自然数 n 证明性质 $P(n)$，只须证明 $P(1)$ 并证明性质 $P(n)$ 蕴涵性质 $P(n+1)$.

当戴德金于 1872 年付印他关于无理数定义(附录 2)的小书时，他还是没有感到自然数公理表述的需要．稍后在同他关于无穷集概念的研究相联系时，他才想到这一点，对此我们将在 §3 中比较详细地考察．他在那里建立了自然数的一系列性质，而皮亚诺于 1889 年把它们作为一组公理，现称为*皮亚诺公理*.

他的表述与帕施和希尔伯特对于几何学的表述(第三章, §4)相同:本原对象是一个集合 $\mathbf{N}^*$,一个元素 $1\in\mathbf{N}^*$以及一个映射 s: $\mathbf{N}^*\to\mathbf{N}^*$,它们满足下面3条公理:

Ⅰ) s 是单射,换言之,若 $s(a)=s(b)$,则有 $a=b$;

Ⅱ) 对每个 a, $s(a)\neq 1$;

Ⅲ) 如果 E 是集合 $\mathbf{N}^*$ 的一个子集,满足 $1\in E$ 和 $s(E)\subset E$,则 $E=\mathbf{N}^*$.

当然, $s(n)$就是一旦定义加法后写为 $n+1$ 的元素. 公理Ⅲ)等价于递推原理(取 E 为使得$P(n)$成立的 $n\in\mathbf{N}^*$构成的集合). 以戴德金为榜样,皮亚诺证明他陈述的公理蕴涵欧几里得和格拉斯曼关于 $\mathbf{N}^*$ 的元素所叙述的全部性质. 他还证明了两个结果,它们可以说证实了数学家给予自然数及其性质的唯一地位:

1) 如果($\mathbf{N}'^*$, $1'$, s')是满足公理Ⅰ),Ⅱ),Ⅲ)的另一本原对象组,则存在一个且只有一个一一映射 f: $\mathbf{N}'^*\to\mathbf{N}'^*$,使得

$$f(1)=1',\ s'(f(n))=f(s(n)).$$

如果把皮亚诺公理看作定义一个结构(第五章, §3),那么可以说如果两个结构满足这些公理,则存在其中之一到另一个的唯一的同构;或者用不大精确的说法,只有一个自然数集.

2) 所述三条公理中任何一条都不是其余两条的推论,其证明通过作出满足这些公理中的两条而不满足第三条的3个本原对象来实现[①].

---

① 对于满足Ⅰ),Ⅱ)但不满足Ⅲ)的例子,只须取集合 $\mathbf{N}^*$为本原集,但令 $s(n)=n+2$. 对于满足Ⅰ),Ⅲ)但不满足Ⅱ)的例子,取由三个元素构成的集合 $E=\{1,a,b\}$为本原集,并令 $s(1)=a$, $s(a)=b$, $s(b)=1$. 最后,对于满足Ⅱ),Ⅲ)但不满足Ⅰ)的例子,取同一集合 E,但令 $s(1)=a$, $s(a)=b$, $s(b)=a$.

[227]

我们注意到从皮亚诺公理推出的自然数的一个重要性质：自然数的每个非空集合 A 存在关于自然数通常序关系的最小元素①.

3. 无穷集

A. 无穷集与自然数

从词源学来说，"无穷"是"有限"的否定②. 按人们的直觉，有限集就是可以数出其对象的集合；从数学上说，一旦在皮亚诺公理基础上建立了自然数理论(本章§2,C)，有限集 F 就成为满足下述条件的集合：存在 F 到 $\mathbf{N}^*$ 中由满足 $1\leqslant m\leqslant n$ 的数 m 构成的子集 Z_n 上的一个一一映射. 自然数理论表明，对于 $p\neq q$，不可能存在 Z_p 到 Z_q 上的任何一一映射，因而数 n 是唯一确定的，这就是 F 的"元素个数"或"基数". 更不容置疑的是：不可能存在 $\mathbf{N}^*$ 到某个集合 Z_n 上的一一映射，换言之，在自然数理论中，$\mathbf{N}^*$ 是无穷集.

有可能不用皮亚诺公理来构造有限集理论：我们只须说，F 是有限的，如果不存在 F 到任一元素 $a\in F$ 的补集 $F\setminus\{a\}$ 上的一一映射. 然后就可使用这一定义与"朴素"集合论的运算——并、补集和积集(第五章，§3,B)——来展开欧几里得算术而不必提及无穷集.

反之，欧几里得几何学必须含有直线上点的集合为无穷的假定，因为两个点之间总有另一个点，又例如取线段中点的运算可以无限地重复进行. 由于亚里士多德不承认无穷集(或他所说

① 只须考虑 A 是有限的情形，因为如果 $a\in A$，我们只须证明满足 $n\leqslant a$ 的自然数 $n\in A$ 的集合 A' 具有最小元素；然后可对 A' 元素的个数进行归纳论证.

② 此处不考虑"无穷"一词有时具有的"没有限度"这个意义.

[228] 的"实无穷"①)的存在性,他找到了一条出路,即声称"线不是由点合成",换言之,不能谈论直线上点的集合!鉴于亚里士多德的声誉,可以指望发布这个敕令只是他的门徒所为,因为这是思维混乱的一个很好例子:按照他的"论证"(《物理学》,Ⅵ,1,$231^a21—232^a22$②),一个点应当同由该点所界的半直线上的点"相接触",但这是不可能的,因为它只能同该点的一个"部分""相接触",而点"没有部分"!!这个所谓的"接触"完全是作者捏造的,它并未出现在任何公理中;他把线同实在对象混淆在一起.

声称证实了这个关于"实无穷"禁律的比较严肃的论证或许可以在伽利略的论著中找到:如果我们把每个自然数 n 与其平方 n^2 配对,就定义了(用我们的语言)$\mathbf{N}^*$ 到 $\mathbf{N}^*$ 的一个真子集上的一个一一映射.但是,伽利略写道,如果自然数平方与自然数"一样多",那就会违反"整体大于部分"的公理;因此我们不能说自然数构成一个集合!令人惊讶的是,柯西仍然赞许地援引上述论证.但在 1851 年,出现了波尔查诺的遗著《无穷的悖论》.这一论著的主要意图是阐述作者的哲学观点;可是在其中我们发现戴德金和康托尔之前所作出的包含与等势之间的清楚区分:他谈到直线 $\mathbf{R}$ 的区间$[0,5]$到区间$[0,12]$上的一一映射即映射 $x \longmapsto 12x/5$,即使区间$[0,5]$是$[0,12]$的一个真子集.

关于无穷集的最早的系统研究归于戴德金和康托尔,两人于 1872 年前后开始,很大程度上同时进行,尽管戴德金阐述他的发现的著作直到 1888 年才出版.如我们说过的(见第五章,§3,B),正是在这里他汇集了"朴素"集合论的基本运算,目的

---

① 亚里士多德(以及他以后的许多数学家)只承认"潜无穷",它意味着对于具有某个性质的对象,在得到有限个这样的对象之后,总还有另一个;而"实无穷"就是这些对象的"全体".

② 中译本:《物理学》,商务印书馆,1991.第 162—165 页.——译注

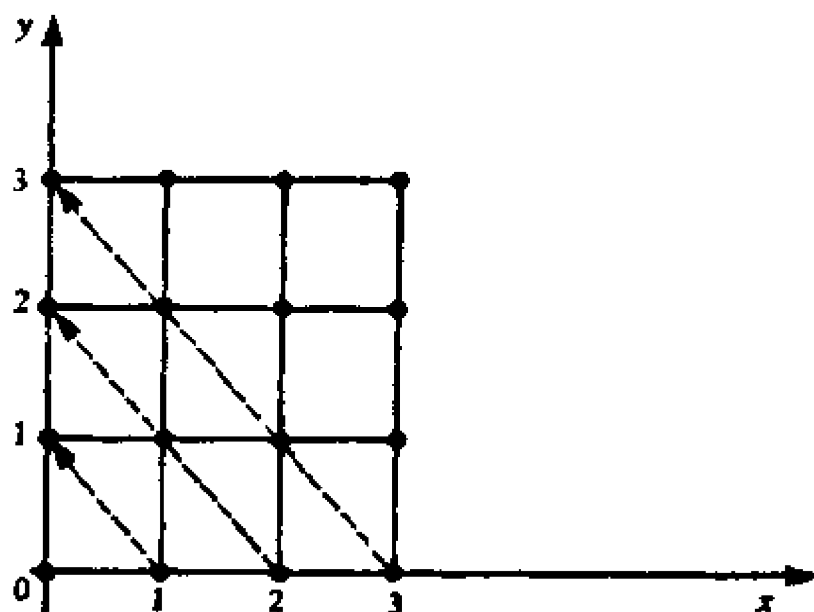
在于澄清无穷集概念与自然数概念之间的关系,这是该书的主要论题.他首次给出的无穷集的数学定义,恰好是在伽利略和波尔查诺的例子中出现过的:集合 E 称为无穷的,如果存在 E 到 E 中异于 E 的一个子集 A 上的一一映射.然后戴德金不假定自然数的存在性证明:任一无穷集 E 必定包含一个子集 $\mathbf{N}^*$,它有一个元素 $1 \in \mathbf{N}^*$,并且存在满足皮亚诺公理(如我们所说,它们是由戴德金首次陈述的)的一个映射 $s: \mathbf{N}^* \to \mathbf{N}^*$.他的富有独创性的证明的原理有如下述:按照假定,存在一个元素 $a \in E$ 和 E [229] 到其子集 $f(E)$ 上的一个一一映射 f,且 $f(E)$ 不含有 a.对于 E 的含有 a 且满足 $f(K) \subset K$ 的子集 K,戴德金把它称为(关于 f 的)链(由于 $f(K) \subset f(E)$,所以 $f(K)$ 不含有 a).集合 $\mathbf{N}^*$ 就是所有链之交,如果我们把 a 写为 1,那么 f 在 $\mathbf{N}^*$ 上的限制 s 和元素 1 满足皮亚诺公理.我们不来给出戴德金证明的细节,也不来给出他通过纯集合论的论证得到自然数所有性质的方法.

B. 无穷集的比较

康托尔对集合论的贡献比戴德金更有独创性.在他开始研究之前,人们只辨认出有限集与无穷集这两种类型的集合,而且也无人试图对无穷集再作什么区分.正是康托尔引起普遍的惊奇,他首次证明存在若干类型的无穷集,它们不能互相归属.

我们已描述过(见第五章,§4,A)康托尔如何引进集合之间的等势关系.首先,他特别在无穷集中区别出可数集,其定义为与自然数集 $\mathbf{N}^*$ 等势的集合.这种集合的重要性来自一个集合的不同元素的序列在分析中所起的作用(如果 (a_n) 是一个序列,则映射 $n \longmapsto a_n$ 可看作 $\mathbf{N}^*$ 到序列元素的集合上的一一映射).康托尔很快看出自然数偶 (m, n) 的集合是可数的:如果我们沿"平行于第二象限分角线的各列"来"数出"这些偶,那么就容易使他的步骤"形象化"(图 53).

这里以考虑集合 $\mathbf{N} = \mathbf{N}^* \cup \{0\}$ 来代替 $\mathbf{N}^*$ 更加方便,这就使



[230]

图 53

康托尔得以“数出”序列

$$(0,0),(1,0),(0,1),(2,0),(1,1),(0,2),$$
$$(3,0),(2,1),(1,2),(0,3),\cdots\cdots$$

一个直接的推论就已使人多少有点惊讶,这就是不可约分数 p/q 的集合 $\mathbf{Q}_+^*$ 是可数的①:事实上我们需要做的是按康托尔的步骤列出数偶(p,q),但只保留 p 和 q 不等于零且没有异于 1 的公因数的分数.

1873 年,康托尔提出了一个从未有人想过的问题:实数集 $\mathbf{R}$ 是可数的吗?他的第一个主要发现是这个问题的答案是否定的.他基于区间套原理通过反证法既简单又聪明地发现了这个答案(附录 3).

这个成就鼓舞了康托尔,使他把自己毕生大部分精力用于等势问题.基于有限集的“基数”模型,他称两个集合具有“相同

① 看来任何两个有理数之间总有无穷多个有理数这个事实阻碍了人们“枚举出”所有有理数.

的势"或"相同的基数",如果它们等势;他进而倾注于在基数间定义序关系.如果存在集合 E 到集合 F 的一个子集上的一一映射,那么自然说 E 的基数"小于或等于"F 的基数.对于有限基数(即自然数),这事实上就是通常的关系 $m\leqslant n$;而每个有限集的基数确实小于(且不等于)可数集的基数,后者以希伯来字母记作$\aleph_0$(读作"阿列夫零").但在我们能够谈论全序关系之前,我们必须先验证刻画这种关系的 4 条公理(第五章,§3,D).自反性显见,传递性也是简单的:如果存在 E 到 F 的一个子集上的一一映射 f 与 F 到 G 的一个子集上的一一映射 g,则合成映射 $g\circ f$ 是 E 到 G 的一个子集上的一一映射.然而在另外两条公理上,康托尔遇到了困难.直到 1898 年,施罗德和 F·伯恩斯坦才证明,如果存在集合 E 到集合 F 的一个子集上的一一映射,也存在 F 到 E 的一个子集上的一一映射,那么 E 与 F 是等势的①(附录 3).至于最后一条公理即对任何两个集合 E,F,总是 [231] *或者*存在 E 到 F 的一个子集上的一一映射,*或者*存在 F 到 E 的一个子集上的一一映射,只有通过在集合论中引进我们以后要谈到的(本章 §4,B)称为策梅鲁公理的公理,才能加以证明.

康托尔及其继任者也发展了无穷基数的"算术",它与通常算术明显不同.任务在于从给定基数的一些集合出发,计算通过集合论中通常运算所得到的集合的基数;这些运算是:没有公共元素的一些集合之并,一些集合之积,以及一个集合的子集构成的集合.如果 $\mathfrak{m},\mathfrak{n}$ 分别是集合 E,F 的基数,则我们分别以 $\mathfrak{m}+\mathfrak{n},\mathfrak{m}\cdot\mathfrak{n},2^{\mathfrak{m}}$ 记集合 $E\cup F,E\times F,\mathfrak{P}(E)$的基数;当 $\mathfrak{m},\mathfrak{n}$ 是无穷基数时,我们有

$$\mathfrak{m}+\mathfrak{n}=\mathfrak{m}\cdot\mathfrak{n}=\sup(\mathfrak{m},\mathfrak{n}), \qquad (1)$$

其中 $\sup(\mathfrak{m},\mathfrak{n})$记两个基数 $\mathfrak{m},\mathfrak{n}$ 中较大的基数.此外,对任何

① 在戴德金的论文中发现了这个定理的证明,其日期可追溯到 1887 年.

基数 $\mathfrak{m}$,有

$$\mathfrak{m} < 2^{\mathfrak{m}}. \tag{2}$$

后一关系表明,如果我们依次考虑集合 $\mathbf{N}$,$\mathfrak{P}(\mathbf{N})$,$\mathfrak{P}(\mathfrak{P}(\mathbf{N}))$,$\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\mathbf{N})))$,……,则我们得到一个无穷序列,其基数全不相同.但是关系式(1)表明,以通常算术为模型来推想真正的"关于无穷的算术",会是毫无意义的.

对于无穷基数 $\mathfrak{m}$,(1)的特殊情形 $\mathfrak{m}^2=\mathfrak{m}$ 当 $\mathfrak{m}=\aleph_0$ 时又给出了康托尔最早关于 $\mathbf{N}\times\mathbf{N}$ 的结果.在对于任意无穷基数证明上述结论之前,康托尔于 1877 年对实数集 $\mathbf{R}$ 的基数 $\mathfrak{c}$ 证明了这个结论:集合 $\mathbf{R}$ 与 $\mathbf{R}^2$ 是等势的.这是康托尔提出的最令人惊奇甚至在当时数学家中引起混乱的定理.从古希腊人以来,一直有这样的信念,即在 1,2,3"维"几何对象(曲线、曲面和空间区域)之间有着深刻的区别,而康托尔的结论像是消除了这种差别,从而摧毁了整个几何学!是戴德金看出如何使康托尔的定理同经典数学协调起来.维数概念不仅是关于 $\mathbf{R}^n$ 的基数的,也是关于此空间的拓扑(第五章,§3,E)的;康托尔定义的 $\mathbf{R}\to\mathbf{R}^n$ 的一一映射是不连续的,而戴德金猜想,当 $m\neq n$ 时,不存在成为同胚(第五章,§4,A)的一一映射 $f:\mathbf{R}^m\to\mathbf{R}^n$.但是一直没有[232]发现这一猜想对于满足 $m\neq n$ 的所有数偶(m,n)成立的证明;直到 1911 年,才由 L·E·J·布劳威尔运用由庞加莱和他本人引进的全新概念给出证明.

对于许多结构,有可能以一般的方法定义维数概念(甚至定义几种维数概念),即与所给结构相联结且关于同构(第五章,§4,B)不变的实数.不仅在拓扑学和同调代数(第五章,§5,A,Ⅷ,Ⅸ,Ⅹ)——在这些分支中能有几种给出不同维数的定义——中是如此,而且在代数几何学(上引ⅩⅩ)和测度论(上引Ⅻ)中也是如此.在最后的情形中,"维数"能取不是自然数的数(甚至是无理数),而且最近它在某些称为"分形"结构的应用中取得重要地位.当一个集合赋予若干结构时,不再有理由指望与

这些结构相联结的"维数"都相同.

4. "悖论"及其后果

A. 存在与构造

"存在"和"不存在"是数学家最常用的词汇之一.从数学对象本性的角度看,当我们说到具有某种性质 P 的数学对象"存在"时,很清楚我们并不意味着它与说到经验世界的某一事物"存在"是同一回事.对于后一情形,"存在性"的证明在于做一个实验,使得问题中涉及的对象可为我们的感官所感知,或者能使我们通过感官可以感知的逻辑推论来推断这种"存在性"——此时它就意味着介于对象与人们感觉之间的一种理论.

从帕施和希尔伯特引进的公理论——这是当代绝大多数数学家采纳的理论——的观点看,"存在性"概念与公理系统有关.它是关于借助所给理论的本原关系表述的性质 P 的问题,并涉及在该理论中本原的或被定义的对象;于是就在该理论中定义了具有性质 P 的对象的集合 E,而说具有性质 P 的对象"存在"就意味着陈述"E 为非空"是该理论中由公理推出的定理①.我们方才(本章 §3,B)就看到实数论中的两个例子:当 $m \neq n$ 时 $\mathbf{R}^m$ 到 $\mathbf{R}^n$ 上的一一映射的集合非空(康托尔)而 $\mathbf{R}^m$ 到 $\mathbf{R}^n$ 上的同胚的集合为空(布劳威尔). [233]

为证明 E 是空集,我们通常用反证法来证明相反的假定在所述理论中导致矛盾.布劳威尔在上述例子中就是用这种方法进行证明的.我们记得,已知的最早用"反证法"的论证(见第三章,§2)正是证明使得 $r^2=2$ 的有理数 r 的集合为空.

反之,具有某种给定性质的对象的存在性则常常通过"构造"来证明.康托尔关于 $\mathbf{R}$ 与 $\mathbf{R}^2$ 等势的定理就是这种情形(附

① 约定当一个结构具有本原对象的集合时,这些集合都不是空的.

录3). 一个更简单的例子是我们描述过的(本章§1,A)欧几里得平面几何学中矩形存在性的证明.

然而也有通过反证法作出的"存在性"证明.最新近而且最引人注目的例子之一是费特-汤普森定理(见第五章,附录2,D).问题是要证明任一奇数阶有限群 G 是可解群.它等价于断言(见前引章节)在 G 中存在异于 G 和 $\{e\}$ 的正规子群.对于这一论断的"构造性"证明就会是从 G 的阶为奇数这个事实出发来"构造"这样的子群;但迄今仍不知道这样做的途径.于是他们通过反证法来论证:假定存在阶为不等于1的奇数的群 G_0,它除自身和 $\{e\}$ 外没有别的正规子群(即是单群);我们甚至可假定这样的群 G_0 具有最小可能的阶数.费特和汤普森的证明在于分析 G_0 的性质,在250页的分析之后,终于得到了一个矛盾.

大多数数学家喜欢存在性的"构造式"证明,它常常提供关于"所构造的"对象的更精确的信息;但当没有别的证明时他们也承认"非构造"证明.

B. 集合概念的变异与选择公理

直到1870年前后,当数学家们谈到对象的集合时,他们指的总是数学对象特别是"经典"对象,诸如数、"图形"、函数等等.大约从1930年起,绝大多数数学家又这样做了.然而其间发生一个有趣的插曲,它使集合论成为数学与哲学的汇合点,其结果是扰乱了许多数学家的心智.这就是所谓的"基础危机".

[234] 首先,康托尔和戴德金这两位无疑要受到责备,因为正是他们觉得有必要来"定义"集合.我们知道康托尔的"定义":"把我们感觉或思维的不同对象收集在一起";戴德金的定义也没有什么差别.这种定义如同欧几里得关于点和线的假定义(见第三章,§4)那样,空洞无物,毫无用处.但他们为这样的想法敞开了门户:"集合"这个词并不只是标志数学对象 E(它同别的数学对象 x 以属于关系 $x\in E$ 联在一起);它也能容纳事物(诸如课文

中的词或句)的总体这一本来意义.

在对这些集合进行论证时,这个时期的数学家都像 17 世纪的数学家论证"无穷小"那样,以自己关于这些集合的独特"直觉"为指引.但是,除康托尔及其继承者在定义基数和某些序关系中的一些混乱外,在一段长时间中这些论证只依赖于"朴素"集合语言(第五章,§3,B),并未引起争辩.

然而,到 19 世纪末叶,数学家们认识到布尔和戴德金容许的集合运算作为分析中许多论证的基础是不完全充分的.例如,设 E 是平面中的点的一个集合,x_0 是平面中不属于 E 的点,但以 x_0 为中心的任一圆盘至少含有 E 的一个点.问题是证明存在 E 的点的一个序列 $a_1,a_2,\cdots,a_n,\cdots$,它以 x_0 为其极限.其论证十分简单:令 $\boldsymbol{D}_n$ 为以 x_0 为中心、以 $1/n$ 为半径的圆盘(图

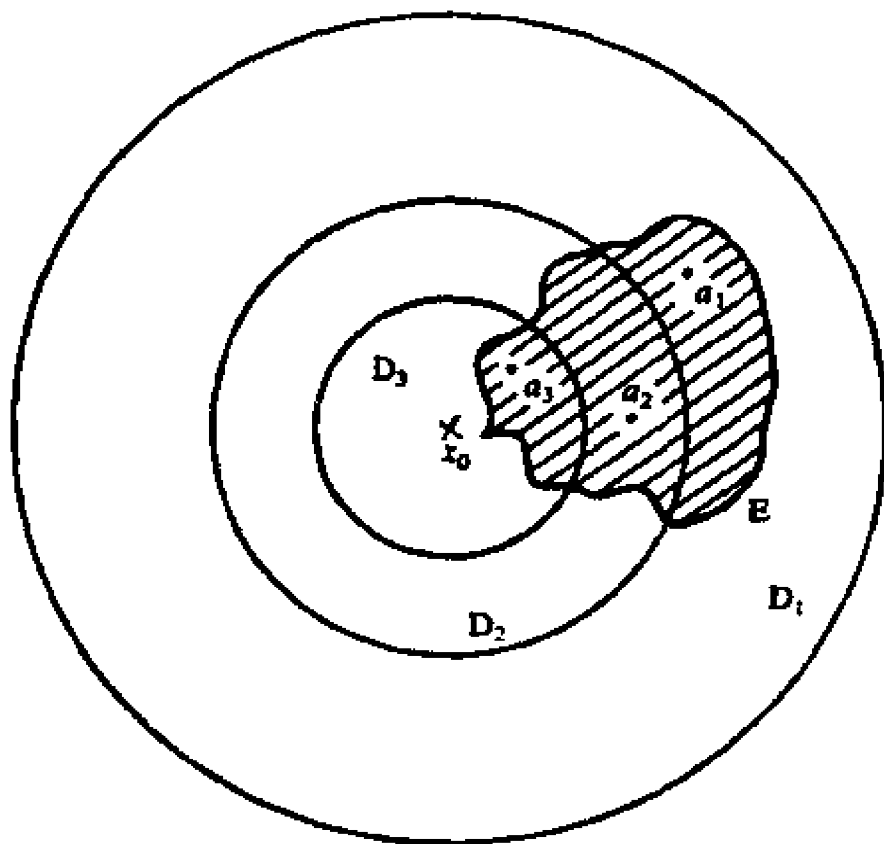


图 54

54). 在圆盘 D_1 中有点 $a_1 \in E$,在圆盘 D_2 中有点 $a_2 \in E$,等等. 这些点构成所求的序列,因为对任一自然数 N,当 $n \geqslant N$ 时,x_0 到 a_n 的距离 $\leqslant 1/N$. 如果推广这个例子独具的性质,那么我们可看出它事实上使用了下述一般原理:设(E_n)是集合 E 的非空子集的一个序列(在前述例子中,$E_n = E \cap D_n$),则存在 $\mathbf{N}^*$ 到 E 中的映射 $n \longmapsto a_n$,使对每个 n 有 $a_n \in E_n$. 看来这个原理是完全自然的结论,在很长时间中它已被使用但甚至都不提到它. 第一个对它加以注意的数学家无疑是皮亚诺,他于 1890 年必须把它用于一个函数集 E. 他提出的要害之点是,对于每个 n,元素 a_n 在 E_n 中无论如何不是唯一确定的(E_n 常是无穷集). 如果"直观地"论证,这就必须在 E_n 中"挑选"a_n,因此就得相继进行无穷多次"挑选",而皮亚诺拒绝这样做. 别的数学家在 20 世纪初追随他的想法,特别是策梅鲁以"选择公理"这个不大适当的名称陈述了一条远为一般并且不太"直观"的原理. 为应用于基数理论,策梅鲁需要下面的命题:如果给定任意一个集合(不一定可数)I 到集合 E 的子集的集合 $\mathfrak{P}(E)$中的一个映射 $U: I \to \mathfrak{P}(E)$,满足:对于每个 $\alpha \in I$, $U(\alpha)$非空,则存在一个映射 f: $I \to E$,使对每个 $\alpha \in I$, $f(\alpha) \in U(\alpha)$. 这个映射就对每个 $\alpha \in I$ "挑选"了 $U(\alpha) \subset E$ 的一个元素. I 为可数的情形就是皮亚诺拒绝的原理,有时称为可数选择公理.

可以看出这种趋势导向何方,这种趋势是赋予集合概念以"实在"特征而不管它涉及的是数学对象. 说集合像是商店中供人"挑选"东西的货架,同说直线是两点之间拉紧的一条线差不多;拒绝策梅鲁公理——它仅是肯定一个数学对象即一个映射的存在性——不比拒绝接受欧几里得第一个"规定"更合理(在大地上不可能伸展从地球到天狼星的直线!).

因此策梅鲁认为,避免无益争吵的最好办法是,像帕施和希尔伯特对几何对象所做的那样(第三章,§4),通过用公理系统来定义集合,使集合概念恢复作为数学对象(在柏拉图意义

下)的特征.这里在任何数学对象之间只有一个“本原”关系 $x \in X$.另一个包含关系 $X \subset Y$ 由上述关系导出,它意味着任何满足 $x \in X$ 的对象 x 也满足 $x \in Y$.然后策梅鲁提出康托尔和 [236] 戴德金的“朴素”集合语言中的运算,并以公理形式陈述它们用到的性质.第一条外延公理已为布尔和戴德金所陈述,即关系“$X \subset Y$ 和 $Y \subset X$”蕴涵 $X = Y$.然后有一系列公理,断言空集 $\varnothing$(对任何对象 x 有 $x \notin \varnothing$)的存在性,偶(一个集合 $\{x, y\}$,它仅有的两个元素是任意两个相同或不同的对象 x, y)的集合的存在性;集合 X 的子集的集合 $\mathfrak{P}(X)$(一个集合,其元素是所有子集 $U \subset X$,因而关系 $U \in \mathfrak{P}(X)$ 等价于 $U \subset X$)的存在性.在这些公理之上他又添加前述的选择公理以及断言无穷集存在性[①]的无穷公理.

C. 悖论与形式化

还有一条最重要的公理尚待陈述,因为正是它把集合引入到数学论证中:它断言任一性质 P 定义一个集合,其元素恰是所有具有性质 P 的对象.这条公理从古典时代以来就已被隐含地使用(例如定义一条“几何轨迹”),它由弗雷格以概括公理的名称明确地陈述.但在把它并进公理式理论之前,由于“性质”概念不精确,因而产生了严重困难.这些困难以集合论中称为“悖论”的形式出现.

第一个悖论以其最简单的形式取 P 为“是一个数学对象”这个性质,这样就会有一个集合 Ω,其元素是所有数学对象.特别是,Ω 会是它本身的一个元素,这多少有点难以想象.更糟糕的是:Ω 的每个子集 U 是一个数学对象,因此是 Ω 的一个元素,于是这些子集的集合 $\mathfrak{P}(\Omega)$ 就会是 Ω 的一个子集,但这与

① 波尔查诺和戴德金认为,通过考虑“思想之集合”这个例子,就有可能证明无穷集的存在性!

断定不存在$\mathfrak{P}(\Omega)$到Ω的一个子集上的一一映射(本章§3,B,公式(2))的康托尔定理矛盾①.

为避开这个矛盾,策梅鲁对他关于概括公理的表述加以限制:他不承认由具有给定性质P的对象构成的集合的存在性,除非这些对象已是早先已定义的一个集合的元素.例如,可以说[237]到偶数的集合,因为所谈的是已定义的集合即自然数集中的对象,而且谈的是该集合中一个元素"被2整除"的性质.

但是,在策梅鲁所做研究的一般性水平上,怎样理解"性质"这个词?策梅鲁限于说"不管一个性质是否有用,它必须由公理和普遍适用的逻辑规则以非任意的方式确定".显然他心中是以直到那时数学家所考虑的性质为典型.但第二个"悖论"表明这不够精密.考虑自然数集$\mathbf{N}^*$,并以下述方式定义自然数的一个"性质"P:考虑英语中所有这样的单词序列,它形成无歧义地定义一个自然数的短语,例如"the largest prime number dividing two hundred and forty - two"("能整除242的最大素数").存在有限个至多具有20个单词的英语短语,因此这种定义一个自然数的英语短语也只有有限个.不能由这些短语之一定义的自然数满足性质P:"not capable of being defined in a phrase of less than twenty words"("不能由少于20个单词的短语定义").

这些数应当构成自然数的一个非空集,因而具有最小的数n.这个数无歧义地由下述短语定义:"the smallest nature number which cannot be defined in a phrase of less than twenty words"("不能由少于20个单词的短语定义的最小自然数");但这个短语只有16个单词,从而我们碰到了矛盾.

尽管这个论证看起来奇怪得像一个谜语,尽管像P这样的"性质"以前从来没有数学家想到过,但有必要表述适应数学家用法的限制,从而避免寄生式"性质"的陈述.弗伦克尔和斯科朗

① 这个论证是称为"罗素悖论"的"悖论"的简化形式.

于 1922 年提出的解决办法在于在数学关系的陈述中消除日常语言,代之以人工语言,它由取自一个固定表列的符号的组合构成,并遵从可以避免通常语言中歧义的硬性文法.这样的语言由数理逻辑的奠基者们于 19 世纪描述(本章, § 5, A),它称为形式语言.在这种语言中,迄今还没有人表述过可以引出"悖论"的性质.

可是,用这种语言表达的数学论证会冗长得吓人.引进了为数众多的简略,使之达到写作当代数学论著的语言.有经验的数学家知道如何辨认可以翻译为形式语言的文词;在出现疑惑之处,由作者引进为使他的同事信服所必须的精密性.

使之这样精密并以形式语言写成的策梅鲁诸公理,构成了 [238] 通常所称的 ZFC 系统.记号 ZF 也用于没有选择公理 AC 的该系统.当今绝大多数数学家使用 ZF 或 ZFC 系统,但通常不明白声明.我们已看到,就这些数学家关注的内容而言,所有数学分支都可归属于集合论,因而最终都可归属于 ZF 或 ZFC 系统.这样的数学家常被称作"形式主义者",尽管他们从来不用真正的形式语言来写作.他们以戴德金和康托尔的"朴素"语言来表述,只要他们一讨论无穷集,就不会对公理的虚假"直观"特性产生任何错觉.关于数学家中的"非形式主义者",我们也会说上几句(本章, § 5, D).

5. 数理逻辑的勃兴

A. 逻辑的形式化

逻辑的应用显然不限于数学,它显示在流传到我们的所有古希腊哲学家的著作中.我们受惠于亚里士多德,因为他开始汇编了逻辑规则,使之系统化,也因为他辨认出这些规则不依赖于他们所讨论的对象或关系的特殊性质.他引进了一个命题或关系的否定,但他主要专心于三段论的分类,而三段论对涵盖逻辑

的全部应用是不够的,在数学中尤其如此.另外两个逻辑运算即两个关系 R 与 S 的合取"R 与 S"以及两者的析取"R 或 S"[①]占着一大部分.

随着代数的发展,不会不注意到逻辑规则与代数运算规则之间的相似性.于是许多人试图创造表示逻辑运算的符号,其中莱布尼茨的符号特别值得注意.但这项工作直到 20 世纪才发表.在布尔及其直接继承者之前,在 19 世纪中叶,无人成功地创造出作为集合语言先导的运算.布尔使一个"全域"的对象的性[239]质与此全域中具有该性质的对象的"类"相联系.如果 x,y 分别是对应于性质 P,Q 的类,则对应于性质"P 与 Q"的类记作 xy(交),而对应于"P 或 Q"的类记作 $x+y$[②](并).我们也需要对于对应于否定"非 P"的类的记号 x'(x 的补).最后,我们以 1 表示全域的所有对象的类,以 0 表示 1 的补即空类.

这就组成通称的"布尔代数",其中有所引进的运算之间的各种恒等式,有些与通常代数中的恒等式相似,例如

$$x+y=y+x,$$

$$xy=yx,$$

$$x(y+z)=xy+xz;$$

有些则出乎意料,例如

$$x+x=x,$$

$$xx=x,$$

$$x+yz=(x+y)(x+z),$$

$$1+x=1.$$

还有"对偶"关系:

---

① 对于后者,与"或"这个字在日常用语中的"析取"意义相反,说"R 或 S"为真意味着或 R 为真,或 S 为真,或两者均为真.

② 起初布尔仅当交 xy 为空时才用符号 +;杰文斯把它推广到用于所有情形.

$$(x')' = x, (x+y)' = x'y', (xy)' = x' + y'.$$

所有这些关系还能从其中数目较少的一些关系推出；通过把后者作为公理，就可认为定义了一个代数结构（见第五章，§5，C）.

对于把数学定理翻译为形式语言，布尔代数同样不甚适用，尤其当定理涉及几个对象之间的关系时更是如此. 例如下面的断言：

对每个正实数 x，存在实数 y，使得 $x = y^2$.

用当代逻辑学家最常用的基本上由弗雷格和皮亚诺创立的形式语言，上述断言翻译为

$$\forall x (x > 0 \Rightarrow \exists y (y \in \mathbf{R} \text{ 与 } x = y^2)).$$

逻辑符号 $\forall$，$\exists$ 称为量词，分别代替“对每个”和“存在”，而 $\Rightarrow$ 意味着“蕴涵”. 经验证明，在 ZFC 系统（本章 §4，C）的规则下，可以用这种方法写出全部数学命题，如果不引进缩写（例如上述例子中的 $>$，0，$\mathbf{R}$，y^2），那么所需要的就只是逻辑符号与数学符号 $=$ 和 $\in$.

B. 元 数 学

19 世纪进程中出现的与经验实在并无联系的形形色色的结构（见第五章，§3），显然为数学家选择结构的公理提供了自由. 但这种自由有一个基本限制，就是康托尔从 1883 年以后一直强调的，所引进的概念必须不是不相容的. [240]

人们怎样才能保证事实上情形确是如此呢？一旦发现能证明非欧几何相容性的模型观念（本章 §1，C），它就毫无困难地应用到许多别的理论上. 实数理论的算术化（本章 §2，A）提供了自然数理论中的实数论模型，从而也使同一理论中的几何学模型有效建立. 至于 19 世纪进程中发明的种种结构，由于它们的引进正是为了对经典理论中提取的例子给出一般的表述，所以它们都具有经典数学中的模型（参见第五章，§2）.

这样,所有事情最终归结为算术的相容性.直到19世纪末,这甚至还是一个无人提出过的问题,有的只是看来自然数是人类心智的最基本概念之一,对于某些人来说,它甚至具有像我们的外部世界那样的“实在性”.

集合论的“悖论”(本章§4,C)结束了这种高枕无忧的状况.这是由于算术的发展,尤其是由于为构造别的理论的“算术”模型,必须把一般集合论的某些公理应用于自然数集(本章附录2),它恰恰具有带来矛盾的危险,并使这些模型变得毫无用处.

正是希尔伯特于1904年后提出和开发了处理相容性问题以及通过更一般的途径先验地检验数学证明能完成什么这个问题的一种新方案,这就是通称的*元数学*或*证明论*.

这一方案依赖于把整个数学(用逻辑符号)表达为形式语言的可能性.形式化数学可与象棋这样的游戏相比较:逻辑符号和数学符号是游戏的棋子,给定理论的公理和定理是游戏规则允许的棋子位置;按照逻辑规则从一条定理转向另一条定理是游戏规则允许的“*走步*”.

于是任务就在于通过抽象出数学中能赋予证明的意义来分析证明,并试图预测它们将在何处终止.这又可与分析棋局相比较,因为在下棋时有可能表明棋局以“将死”告终或是永远不能[241]达到这一结局.

C. 数理逻辑的凯旋

正是在希尔伯特概述的框架中,当代逻辑学家取得了一系列完全出人意料的结果.它们使人们对集合观念有了新的认识,但根本不能说已达到了希尔伯特过分乐观的预期.

希尔伯特在其学派的几位数学家的协助下,曾开始把他的方案应用到比皮亚诺公理系统(本章§2,C)更简单的公理系统上,这样就能不包括自然数的所有性质.对于这些公理,可以证明不会有任何矛盾.

然而,1930年,年青的数学家K·哥德尔发表了一篇短文,他在文中宣布下述两者之一为真:或者算术是不相容的,或者它是相容的但不可能证明其相容性.这样一来,希尔伯特的期望彻底落空.

这个异乎寻常的结论只有通过仔细研究哥德尔的方法才能理解,而他的方法却极其复杂.我们只能对他的原理提供十分粗糙的想法①.

哥德尔的证明沿着两条路线进攻.他从描述一本"词典"开始,通过各方面都详尽、明白、精确的步骤,他先使形式语言中的每个符号与一个自然数相对应,然后再使这些符号的每种为规则所允许的组合与一个自然数相对应,最后使每个依据同样规则组成一个证明的这些组合的序列与一个自然数相对应.

这使他能够描述算术中一个命题 A 的构成,并且证明,如果此算术不是不相容的,则在此算术中不存在 A 和非 A 的证明.我们称 A 在此算术中是不可判定的.希尔伯特认为能证明算术中每个命题在该算术中是"可判定的",而哥德尔定理则通称为算术的不完全性定理.

在他的第二条进攻路线中,哥德尔描述了算术中另一命题 C 的构成,并且仍然借助他的"词典"证明,如果在此算术中存在 C 的证明,则此算术就会不是不相容的.此外他还证明,陈述"C 蕴涵 A"是此算术中的一条定理;如果在此算术中存在 C 的 [242] 证明,那就也会存在 A 的证明,但如果此算术不是不相容的,则由不完全性定理,不可能存在 A 的证明.

① 由于这个结论十分奇特,所以它吸引了许多通俗科学杂志的出版者,在这些杂志上专门发表了许多有关的文章.如果我们加上四色定理、费马方程 $x^n+y^n=z^n$ 以及电子计算机所用的逻辑规则,那我们就已为这些期刊上出现的数学问题开出了一张相当完全的清单.他们设想这些期刊的读者相信,除此之外没有什么别的问题了!

哥德尔描述的不可判定命题 A 显得非常人工斧凿,与当前数论中任何其他部分都没有任何联系.它的主要功用是证明算术相容性的不可能性.在数论中许许多多尚未解决的经典问题(见第四章)中,据我所知,其中没有一个已被证明是不可判定的.

另一方面,哥德尔本人于1940年以及随后P·J·科恩于1963年所完成的却真可以说是集合论中的一场革命.康托尔作出了一个猜想,不知如何去证明,即在实数集 $\mathbf{R}$ 中,每个无穷子集或者是可数的,或者与 $\mathbf{R}$ 是等势的;这称为连续统假设(CH).许多人试图证明CH或非CH;但哥德尔和科恩却发现,如果接受ZF不可能导致矛盾,那么在基于ZF公理的集合论 $\mathscr{S}$ 中,CH是不可判定的.

上述论断的证明是极其专门性的.他们的证明仍然基于模型观念,但却用了新的方法.哥德尔定义了 $\mathscr{S}$ 的一个对象 x 的一种性质 $R(x)$,它表达了能以某种详尽描述的途径"可构造"这个事实;而 $\mathscr{S}$ 的可构造对象就是使得 $R(x)$ 是 $\mathscr{S}$ 的一条定理的 x.理论 $\mathscr{S}'$ 以 $\mathscr{S}$ 的可构造对象为其对象,并以 $x \in X$ 作为其本原关系,但赋以 x 和 X 两者都必须为可构造这个条件.由这个关系能验证ZF的公理仍然是 $\mathscr{S}$ 的定理,如果它们以同样方式限制到 $\mathscr{S}'$ 的对象上.例如,如果 X 是 $\mathscr{S}'$ 的一个集合,则 X 的又是 $\mathscr{S}'$ 的集合的子集是 $\mathscr{S}'$ 的一个集合的元素.如果把ZF的公理的这些"限制"取作 $\mathscr{S}'$ 的公理,则这样定义的 $\mathscr{S}'$ 是 $\mathscr{S}$ 的一个模型,并且通过限制属于 $\mathscr{S}'$ 的对象为可构造对象,$\mathscr{S}$ 的每个关系可"翻译"到模型 $\mathscr{S}'$ 中.

这样做后,令CH′是连续统假设CH在 $\mathscr{S}'$ 中的"翻译";哥德尔选取的关系 $R(x)$ 是CH′是 $\mathscr{S}'$ 的一条定理,因此CH′也是 $\mathscr{S}$ 的一条定理.于是如果"非CH"是 $\mathscr{S}$ 的一条定理,则通过"翻译","非CH′"就会是 $\mathscr{S}'$ 的一条定理,因而也是 $\mathscr{S}$ 的一条定理,又因CH′也是 $\mathscr{S}$ 的一条定理,所以 $\mathscr{S}$ 就会是不相容的.我们必

须细心地注意，这表明 $\mathscr{S}$ 中没有"非 CH"的证明，但并没有证明 $\mathscr{S}$ 中有 CH 的证明. [243]

1963 年，P·J·科恩发明了构造 $\mathscr{S}$ 的模型的一种新的特别灵活的方法，称为"力迫法". 我们从哥德尔的模型 $\mathscr{S}'$ 开始，对 $\mathscr{S}'$ 的对象添加 $\mathscr{S}$ 的新对象. 这些对象是理论 $\mathscr{S}''$ 的本原对象，而这个理论的本原关系 $x \in X$ 仍如前述由"限制"定义. 但这时通过适当定义 $\mathscr{S}''$ 的对象(这是科恩方法的关键之处)，就可能导致：如果 CH″是 CH 在 $\mathscr{S}''$ 中的翻译，则"非 CH″"是 $\mathscr{S}''$ 的一条定理，从而是 $\mathscr{S}$ 的一条定理. 如同前面那样，我们由此作出结论：$\mathscr{S}$ 中没有 CH 的证明. 其结果是，假定 $\mathscr{S}$ 不是不相容的，则 CH 在 $\mathscr{S}$ 中是不可判定的.

现在我们注意到，一般地说，如果 P 是 $\mathscr{S}$ 中的一个不可判定命题，那么我们能考虑在策梅鲁列举的公理上添加 P 或"非 P"所得到的公理 ZF + P 或 ZF + (非 P). 如果 ZF 不导致矛盾，那么这一点对 ZF + P 与 ZF + (非 P)也同样为真. 这样我们定义了两种"集合论"，当然两者互不相容，但都如同 $\mathscr{S}$ 一样成立. 但是还有更多的结论来临：仍把连续统假设表达为没有满足 $\aleph_0 < \mathfrak{m} < c$ 的基数 $\mathfrak{m}$. 对每个整数 $n \in \mathbf{N}$，假定 CH_n 表述为：恰有 n 个不同的基数，满足上述不等式(在这种记号下，CH 就与 CH_0 相同). 于是通过对 $\mathscr{S}'$ 添加适当选取的本原对象来得到模型 $\mathscr{S}_n$[①]，用科恩的方法就能证明，如果 ZF 不导致矛盾，那么 CH_n 在 $\mathscr{S}$ 中是不可判定的，并且这推出 ZF + CH_n 也不导致矛盾. 这样我们就有无穷多种"集合论"，其中任何两个都互不相容！这还不是科恩方法所能做的全部：在最近 20 年中，逻辑学家能证明 $\mathscr{S}$ 中不少命题是不可判定的，特别是，选择公理 AC 甚至可数

① $\mathscr{S}'$ 的两个集合可以在 $\mathscr{S}'$ 中不等势但在 $\mathscr{S}_n$ 中变为等势，因为 $\mathscr{S}'$ 的集合上添加的新集合可以满足：它们之间允许定义一个在 $\mathscr{S}'$ 中不存在的一一映射.

选择公理都是不可判定的.

D. 数学家的反应

由于ZF公理系统蕴涵关于算术的皮亚诺公理(本章§3,A),所以哥德尔不完全性定理表明,不可能通过元数学步骤来建立皮亚诺公理不会导致矛盾.称作——其实并不正确——"形[244]式主义者"的数学家(今天这差不多意味着做研究的数学家的整个群体)把希望寄托在这一点上.就像物理学家和生物学家相信自然定律的永恒性只是由于它与迄今为止他们的观察相符那样,这些数学家相信集合论中不会显现矛盾,因为在最近80年中没有出现过矛盾.再则,这些数学家还相信,设想我们会有无穷集的"直觉"概念是徒劳的;事实上,我们关于小的有限集的不容置疑的直觉使得关于无穷集的不少陈述显得是"自然的",而这些陈述从ZF公理看仍是不可判定的.从同样的直觉观点看,ZF公理也是"自然的".

在反对这种信赖中出现了一批数学家,他们不满意于这样的观点:不允许谈论"真确性",但严格坚持按照公理系统"能证明"的概念①.这些异议者从来数目不多,但却包括诸如H·庞加莱,E·波莱尔,勒贝格和H·外尔这样的著名人物.其中最系统的是L·E·J·布劳威尔,他开始时由于在代数拓扑中做了第一流工作而引人注目,但却把自己大部分科学生涯奉献给重新构筑符合他的想法的数学,并建立了称为"直觉主义者"的学派.还有另一些数学家,其观点接近直觉主义者,称为"构造主义者".

要描述这些数学家的观念是相当困难的;无论如何,他们相互之间也不尽相同.他们承认数学对象具有不同于经验实在的

---

① 所谓"形式主义者"数学家并不禁止谈论"真命题",但对他们而言,这总是指一个命题具有基于明白陈述的公理系统的已发表的证明这种情形;而"假命题"则是其否定在前述意义下为"真"的命题.

"实在性"(或许类似于柏拉图赋予他的"理念"的"实在性"?).他们一致同意自然数序列这种例外情形,对此他们认为他们具有一种基本的直觉,而且他们确信它不会含有隐藏的矛盾.他们不接受一个数学对象的"存在性",除非他们能把它"构造"出来(本章 § 4,A),这意味着大于可数的基数的集合对他们而言只不过是数学范围以外的"幻想".这导致他们限制经典逻辑的规则,并拒绝使用所谓"排中律",它断言对任何命题 P,命题(P 或(非 P))总为真,即使当 P 或(非 P)都不能证明时也是如此.这种态度能使他们避免"悖论",但"形式主义者"接受的许多定理,"直觉主义者"或"构造主义者"却不接受. [245]

两个阵营之间在 1925 至 1930 年期间争论得非常激烈,每一方力求说服另一方,但徒劳无功.时下某些"构造主义者"热衷于传播其信仰,但在数学界几乎没有什么反响.被称为"形式主义者"的年青数学家甚至都不知道曾经有过"基础危机".当我们对数学中发生的事情与物理学中的现实危机加以比较(在物理学中相对论和量子力学迫使物理学家根本改变他们关于自然现象的概念)时,无论如何要说这里"危机"是一个完全误导的词汇.在数学中,我们至多能说"悖论"引起的某种不安;但除布劳威尔及其门徒外,它几乎并未导致什么数学家对他们表达证明的方式作些微改变.更有甚者,据称发生"危机"的 1895 至 1930 年,是数学历史上最丰富多产的阶段之一,第五章 § 5,A 中列举的几乎所有分支在此期间都取得了决定性的进展.在当代,我们亲眼目睹前所未有的古老问题的解决和崭新方法的发现,再说什么"危机"就更加毫无意义了.

E. 数学与逻辑之间的关系

从 1930 年以来,数理逻辑专家并未把他们的努力限于证明集合论中命题的不可判定性.他们也研究许多有别于经典逻辑(它并入到 ZF 系统中)的逻辑体系;不仅逻辑规则可为直觉主

义者接受,而且还有别的逻辑体系,其中有些导源于古代哲学家.这里有模态逻辑、多值逻辑①等等.他们同样彻底分析了数学家们所用的"构造"步骤(本章 §4,A),创造了称为"递归函数"的理论.研究这些问题的逻辑学家的人数不断增加,从而现[246]今数理逻辑已成为一座令人印象深刻的宏大建筑.

某些在其证明中使用经典逻辑的数学家,借助逻辑学家的发现来提供尚未解决的问题的(肯定或否定)答案,关于一般拓扑、有序集理论、测度论以及代数学中某些部分的问题尤其如此.从 20 世纪初叶起,他们为此目的应用了选择公理和连续统假设,而现在他们能通过其公理系统为 ZF + P 的理论来探讨这些问题,这里 P 是 ZF 公理系统不可判定的许多命题之一.

也已注意到,某些新的逻辑系统能翻译为代数结构,恰如经典逻辑翻译为"布尔代数"(本章 §5,A)那样;而这些结构,除它们对于逻辑的意义外,也因其自身的存在价值而得到研究.

同样,如果人们仔细查阅《数学评论》,就能注意到这些工作只代表"形式主义者"所做研究的一个极小部分.大部分工作是在集合的基数几乎都不大于 **R** 的基数以及除可数选择公理(本章 §4,B)从无必要应用 ZF 之外的公理的各个领域(参见第五章,§5,A)中做的②.在这些领域中工作的数学家很少知道逻辑学家在做些什么,如果他们有时听到逻辑学家所做的工作,也不比对于离数学较远的学科诸如生物学或地质学赋予更多的关注.如果逻辑学家有时对这种疏远状况感到惊讶,那是因为他们未能考虑最近 50 多年中数学的演进.逻辑学家没有做任何有助

① 除对经典逻辑的命题赋予"真"值或"假"值外,这些逻辑还允许第三个"可能"值,甚至允许许多其他的值.

② 这些数学家中有一些走得更远,在其论著中宣布他们对选择公理"深恶痛绝".他们的态度与 1930 年的"青年造反派"正好相反,后者为使传统主义者震动,甚至在可以不用选择公理时也要使用选择公理.

于数学家解决他们的问题的事情,而这些问题才是数学家真正感兴趣的;因为如果数学家们偶尔听到什么"悖论",就他们所关心的事情而言,这只不过是一些假问题而已.

6. "严格证明"的概念

各个时代都有数学家批评他们的前辈或同时代人的证明 [247] "不够严格";而他们提供的代替有缺陷证明的论证转过来又常常被他们的后代认为不够适当.这种表面上不断使数学家的论证成为问题的状况最终导致某些人设想,这些论证包含根本的缺陷,这表明永远不可能找到对任何批评都无懈可击的证明.

我认为由于对数学史的误解,我们在这里听到了一种肤浅的看法.让我们再来看一下表达为一系列推断的证明展开的途径:为证明陈述 Q,我们从一个先前的陈述 P 已得证明(或作为一个公理)这个事实出发来证明"P 蕴涵 Q".由于只当构成一个证明的每个推断都正确时才能说该证明为"严格",因此有必要考察什么东西能使得一个推断不正确.

Ⅰ) 撇开平庸的计算错误不谈,数学家可能无意地把 P(对应地,Q)与类似的命题 P'(对应地,Q')混淆起来,而 P' 以及"P' 蕴涵 Q"(对应地,P 以及"P 蕴涵 Q'")已被证明.当 P 的陈述很长、很复杂或当 Q 的陈述涉及多种情形,它们必须分别验证时,常发生这种情形.有时发现这类错误得等上好几十年.

Ⅱ) 命题 P 或"P 蕴涵 Q"不是公理,也未被证明,但显得十分合乎情理.这属于不充分性范畴;我们曾指出欧几里得《几何原本》中的若干例子(见第三章, §4).某些类似的例子是 19 世纪前期的分析学家诸如柯西、阿贝尔甚至狄里克雷、黎曼在连续性和收敛性问题上所犯错误的根源.他们所研究的对象有正确的定义,但他们未能逐步检查,确保他们并未使用这些定义以外的任何东西.类似的情形在 1940 年之前的代数曲面理论中也常出现,就是有意识地忽略一般论证不能应用的"例外情形".

这类错误在已经长期建立的理论中不大普遍.在分析中,从魏尔斯特拉斯时代以来,它们并未出现过.然而即使在今天,也会发生一个证明涉及大量步骤,其中某些步骤的确立需要大量努力;反过来,陈述"P 蕴涵 Q"也可能显得简单,常常多少类似于标准的论证方式.于是就会倾向于一略而过,什么也不写.为过多细节所苦的数学家,就会免除书面证明,而代之以一个老是 [248] 出现的短语"易见……".我不相信会有许多数学家在其一生经历中一次也没有碰到过这样的灾难.

Ⅲ) 在前述各种情形中,为矫正证明需花大量时间是罕见的.当所讨论的对象未被精确定义,因而号称证明的论证从一开头就有问题这种情形则相当不同.在分析中,这种情形发生于17,18世纪关于"无穷小"与级数之"和"的论证中,当时还没有人能说清楚这两个概念意味着什么.当然,在绝大多数情形,那个时代最好的数学家关于对这些模糊概念能做些什么有着正确的观念,而且他们的卓越发现鼓舞他们加速前进,但是他们未能以恰当的数学语言表达他们的证明.在19世纪,虽然不难对他们的结论提供完全正确的证明,但这要等到作为基本概念的极限观念制定出来而且其性质得到全面汇编(见第三章§9)后才有可能.

在整个19世纪,关于拓扑概念(见第五章§3,E)的详细阐述,再度发生同样的情形.具有洞察天才的黎曼创造了数学地研究曲面的拓扑的方法,例如,通过研究曲面上的曲线,使得球面与环面之间的"直观"差别有可能合理化.然而黎曼及其同时代人都没有这种研究必不可少的工具,因而当他们谈到"曲线"、"曲面"或"形变"时,都不得不转而依靠"直观",而这样做的危险联系到皮亚诺"曲线"(本章§2,B)时就很快显现了出来.

然而即使在被看作属于"初等几何学"的"多面体"情形,甚至也有更糟的事情.在其《几何原本》中,欧几里得只定义了棱柱、棱锥和正多面体.1750年,欧拉谈到"多面体",但未给定义.他还说,如果 s,a,f 分别是多面体 P 的顶点、棱和面的数目,则数

$$\chi(P)=s-a+f$$

总是等于 2. 他的证明(无论如何是不充分的)使人想到他涉及的只是凸多面体,即位于其每个面的一侧的多面体. 在 19 世纪进程中,不仅欧拉关于凸多面体的定理得到了正确证明,而且有十来位数学家关注对所有多面体计算数 $\chi(P)$. 问题在于,据我所知,其中没有一个人能给出应当怎样理解"多面体"的一般定 [249] 义. 只是在庞加莱 1895 年的论著中才出现了这样的定义①. 先前这些数学家中每一位都设想他给出了这样一种定义,符合他个人关于多面体应当怎么样的"直观". 稍后另一位数学家显示了一些"图形",他认为也是"多面体",但并不满足他的前人设定的条件. $\chi(P)$的值自然依赖于所选用的"定义",以致看来不会有 $\chi(P)$的一般表达式,其推导是充分"严格"的!

结论是显然的:除非在公理化理论的体系中,否则就不会有"严格"的证明. 在公理化理论中,指明了"本原的"对象和关系,毫无遗漏地列举了这些对象和关系据以联系的公理;因而如果不算Ⅰ)和Ⅱ)中提到的疏忽和错误,那么这个必要条件也是充分的."缺乏严格性"仅仅意味着"缺乏精密性".

历史每每证实这一论断. 对算术中的"严格"证明从来没有过争论;在分析中,从魏尔斯特拉斯以来,在代数拓扑中,从1930 年以后,在代数几何中,从 1950 年以后,都没有出现过争论. 当然,未来的数学家并非不可能需要开发一种理论但不把它

---

① 这个定义是简单的,但得好好想一想! $\mathbf{R}^3$ 中的一条开线段是去掉两个端点的一直线段,而一个开多边形是位于 $\mathbf{R}^3$ 的某一平面上的去掉边界的有界凸多边形. 一个多面体是 $\mathbf{R}^3$ 的子集的一个有限集 $\mathscr{F}$ 的并, $\mathscr{F}$ 的元素是点、开线段和开多边形,而 $\mathscr{F}$ 的任何两个不同元素没有公共点,且满足下述条件:如果开线段的端点属于 $\mathscr{F}$,则此开线段属于 $\mathscr{F}$;如果开多边形的顶点和边(看作开线段)属于 $\mathscr{F}$,则此开多边形属于 $\mathscr{F}$. 亚历山大于 1915 年证明,对于多面体 P, $\chi(P)$关于同胚是不变的.

置于公理化形式之中;但除非他们自己或别人能设法做到使之具有公理化的形式,否则他们就冒着被数学界视为“不严格”的[250] 风险.

附　　录

1. 曲面上的几何学

为简略起见,我们称一个或几个实变量的实值函数为正则的,如果它具有直到3阶的导数(偏导数).

A. 挠　曲　线

在微分几何中,空间 $\mathbf{R}^3$ 中的一条正则挠曲线 Γ 是坐标为

$$x=f(t),y=g(t),z=h(t) \tag{1}$$

的点 $M(t)$ 的轨迹,其中 t 在 $\mathbf{R}$ 的区间 I 中变动,而 f,g,h 是该区间上的正则函数.我们称 t 为曲线 Γ 上点 $M(t)$ 的参数,而公式(1)就是 Γ 的参数表示.例如,圆柱螺旋线由参数表示

$$x=a\cos t,y=a\sin t,z=ht \tag{2}$$

定义,其中 t 在 $\mathbf{R}$ 中变动,a,h 是非零常数.

欧拉揭示,Γ 的从曲线上点 A 到终点 $M(t)$ 的弧的长度 $s(t)$ 是 t 的函数,它具有由

$$s'(t)=((f'(t))^2+(g'(t))^2+(h'(t))^2)^{1/2} \tag{3}$$

给出的导数.例如,对于由(2)给出的螺旋线,有 $s'(t)=(a^2+h^2)^{1/2}$,因而

$$s(t)=(a^2+h^2)^{1/2}t+\text{常数}.$$

B. 曲面上的曲线

对于嵌入于 $\mathbf{R}^3$ 中的“正则”曲面 Σ,高斯的想法是 Σ 同样是坐标为

$$x=f(u,v),y=g(u,v),z=h(u,v) \tag{4}$$

的点 $M(u,v)$ 的轨迹，其中 u,v 分别在 $\mathbf{R}$ 的一个区间内变动，f,g,h 是正则函数. 通过取 u,v 为 $\mathbf{R}$ 的一个区间上的 t 的正则函数并把这两个函数代入(4)，就得到一条描绘于 Σ 上的正则曲线 Γ. 对于这样的曲线的长度 $s(t)$，公式(3)给出

$$(s'(t))^2 = E(u'(t))^2 + 2Fu'(t)v'(t) + G(v'(t))^2, \quad (5)$$

其中 E,F,G 是 u,v 的由下式定义的函数：

$$\begin{aligned} E(u,v) &= \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial h}{\partial u}\right)^2, \\ F(u,v) &= \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + \frac{\partial g}{\partial u}\frac{\partial g}{\partial v} + \frac{\partial h}{\partial u}\frac{\partial h}{\partial v}, \\ G(u,v) &= \left(\frac{\partial f}{\partial v}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2 + \left(\frac{\partial h}{\partial v}\right)^2, \end{aligned} \quad (6)$$

[251]

而在 $s'(t)$ 的公式中，E,F,G 中的 u,v 必须以其 t 的表示式代入. 对于函数 $\varphi(t)$ 的导数，用莱布尼茨的记号

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t}$$

是方便的，于是公式(5)可写为

$$\mathrm{d}s^2 = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2, \quad (7)$$

它具有不特别指明所选参数 t 的优点.

如果Σ上两条曲线 Γ_1,Γ_2 通过同一点，则此两曲线在该点处的切线(适当给定方向)之间的夹角 α 由下式给定：

$$\cos\alpha = \frac{E\mathrm{d}u_1\mathrm{d}u_2 + F(\mathrm{d}u_1\mathrm{d}v_2 + \mathrm{d}u_2\mathrm{d}v_1) + G\mathrm{d}v_1\mathrm{d}v_2}{(E\mathrm{d}u_1^2 + 2F\mathrm{d}u_1\mathrm{d}v_1 + G\mathrm{d}v_1^2)^{1/2}(E\mathrm{d}u_2^2 + 2F\mathrm{d}u_2\mathrm{d}v_2 + G\mathrm{d}v_2^2)^{1/2}}. \quad (8)$$

Σ 上的测地线定义为曲线 $v = \varphi(u)$，其中 φ 满足某个只依赖于函数 E,F,G 及其导数的二阶微分方程.

最后，如果 Σ 上的集合 Δ 对应于平面 $\mathbf{R}^2$ 的点 (u,v) 的一个集合 D，则

$$\Delta \text{ 的面积} = \iint_D (EG - F^2)^{1/2}\mathrm{d}u\mathrm{d}v. \quad (9)$$

有必要提醒一下,对于整个曲面 Σ,并不一定总能找到参数表示式(4),使得 Σ 的每个点仅对应一个参数偶 (u,v). 例如,对于以 O 为中心、以 R 为半径的球面,我们有作为经度 φ 和纬度 θ 的函数的下述参数表示式:

$$x = R\cos\theta\cos\varphi,\ y = R\cos\theta\sin\varphi,\ z = R\sin\theta,$$

其中 $0\leqslant\varphi\leqslant 2\pi$, $-\frac{\pi}{2}\leqslant\theta\leqslant\frac{\pi}{2}$. 但 $\varphi=0$ 和 $\varphi=2\pi$ 对应同一点,且 $\theta=\frac{\pi}{2}$ 时对 φ 的任何值都给出球面的北极. 因此有必要只对球面的由

$$0<\varphi<2\pi,\ -\frac{\pi}{2}<\theta<\frac{\pi}{2}$$

定义的部分应用上述公式.

黎曼的观点在于不管曲面在空间中任何形式的"嵌入",把"曲面部分"定义为 $\mathbf{R}$ 的两个开区间 $a<u<b$, $c<v<d$ 之积,但赋予一个曲线"长度"的定义,此定义一般与欧几里得平面上曲线长度的定义不同. 曲线长度由公式(7)给出,但现在只假定 E,F,G 是 u,v 的正则函数,满足 $E>0$, $G>0$, $EG-F^2>0$, 而不是从诸如(6)这样的公式导出 E,F,G.

例: **平坦环面.** 环面是由半径为 r 的圆绕位于圆所在平面且与圆心相距为 $a(>r)$ 的直线旋转得到的旋转曲面. 取旋转轴为 Oz 轴,就有参数表示式

$$x=(a+r\cos u)\cos v,\ y=(a+r\cos u)\sin v,\ z=r\sin u,$$

[252] 其中 $0\leqslant u\leqslant 2\pi$, $0\leqslant v\leqslant 2\pi$.

如果此曲面上的曲线长度由欧拉公式(3)计算,则得

$$\mathrm{d}s^2=(a+r\cos u)^2\mathrm{d}v^2+r^2\mathrm{d}u^2.$$

另一方面,对于"平坦环面",我们简单地取

$$\mathrm{d}s^2=\mathrm{d}u^2+\mathrm{d}v^2.$$

按照黎曼的观点,其结果是由 $0<u<2\pi$, $0<v<2\pi$ 定义的曲

面部分与平面上该子集不加区别.

C. 庞加莱半平面

现在考虑这样的曲面,它是由点(x,y)构成的半平面 H,其中 $x\in\mathbf{R}$ 为任意,而 $y>0$. 取

$$ds^2=\frac{1}{y^2}(dx^2+dy^2).\qquad(10)$$

圆心在 Ox 轴上的半圆由参数表示

$$x=a+r\cos\varphi,\ y=r\sin\varphi$$

给出,其中 $0<\varphi<\pi$,于是由(10)推出此曲线的长度由

$$ds=\frac{d\varphi}{\sin\varphi}\qquad(11)$$

给出,由此得到

$$s=\log\tan\frac{\varphi}{2}+\text{常数}.$$

对于两个半圆之间的"非欧"角 α(见图 55),由(11)从公式(8)得到

$$\cos\alpha=\cos(\varphi_2-\varphi_1).$$

因此"非欧几里得"角 α 等于"欧几里得"角 $\varphi_2-\varphi_1$.

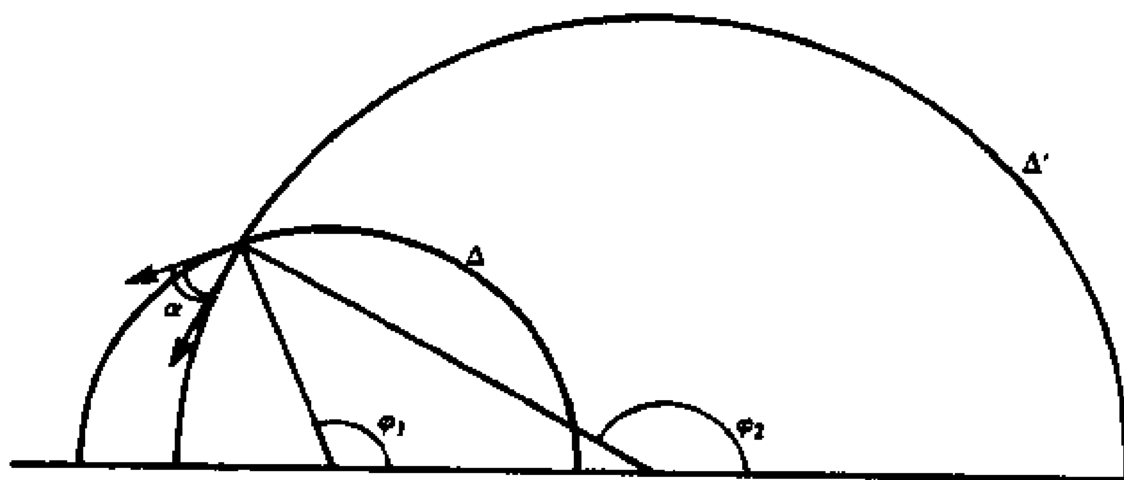


图 55

为研究群 $GL^+(2,\mathbf{R})$的作用,把公式(10)写成

$$ds^2=\frac{1}{y^2}dz\cdot d\bar{z} \tag{12}$$

[253] 的形式是方便的.现在如果

$$Z=X+iY=\frac{\alpha z+\beta}{\gamma z+\delta},$$

则有

$$\bar{Z}=\frac{\alpha\bar{z}+\beta}{\gamma\bar{z}+\delta},$$

因之

$$dZ=\frac{\alpha\delta-\beta\gamma}{(\gamma z+\delta)^2}dz, d\bar{Z}=\frac{\alpha\delta-\beta\gamma}{(\gamma\bar{z}+\delta)^2}d\bar{z},$$

$$Y=\frac{\alpha\delta-\beta\gamma}{|\gamma z+\delta|^2}y.$$

从而

$$\frac{1}{Y^2}dZ\cdot d\bar{Z}=\frac{1}{y^2}dz\cdot d\bar{z}.$$

至于所说群在相同(非欧)长度的"非欧""线段"上的"传递性",我们注意群 $GL^+(2,\mathbf{R})$包含位似和平移

$$z\longmapsto\alpha z+\beta$$

($\alpha>0$,β 是任一实数)的子群.

使用上述子群,只须证明半圆 $x^2+y^2-2x=0$ 上每个弧 $\overset{\frown}{A_1A_2}$能被 $GL^+(2,\mathbf{R})$的一个元素变换为平行于 Oy 的具有相同非欧长度的线段 B_1B_2.变换

$$Z=4-\frac{4}{z}$$

把所说半圆上的点(x,y)变换为点(X,Y),其中 $X=2$, $Y=2y/x$,且有(图 56)

$$\log Y_2-\log Y_1=\log\tan\frac{\varphi_2}{2}-\log\tan\frac{\varphi_1}{2}.$$

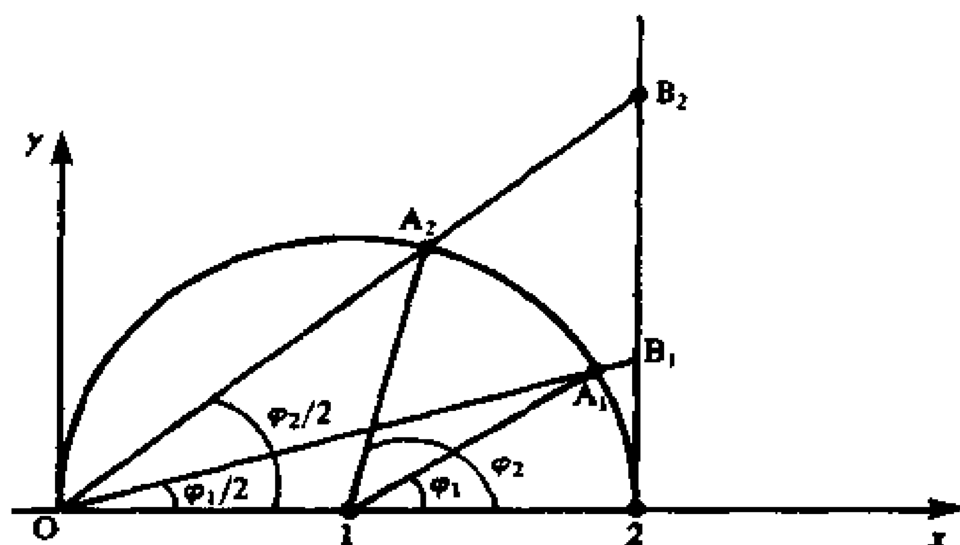


图 56

这一传递性可用来计算三角形 ABC 的"非欧"面积.由于 $\mathrm{GL}^+(2,\mathbf{R})$ 保持角度和面积不变,因此只须对 BC 为 Oy 的平行线的一部分的情形进行计算(图 57).由(9)我们有 [254]

$$
\begin{aligned}
ABC\ 的面积 &= \int_p^q \mathrm{d}x \int_{r\sin\varphi}^{R\sin\psi} \frac{\mathrm{d}y}{y^2} \\
&= \int_p^q \frac{\mathrm{d}x}{r\sin\varphi} - \int_p^q \frac{\mathrm{d}x}{R\sin\psi}.
\end{aligned}
$$

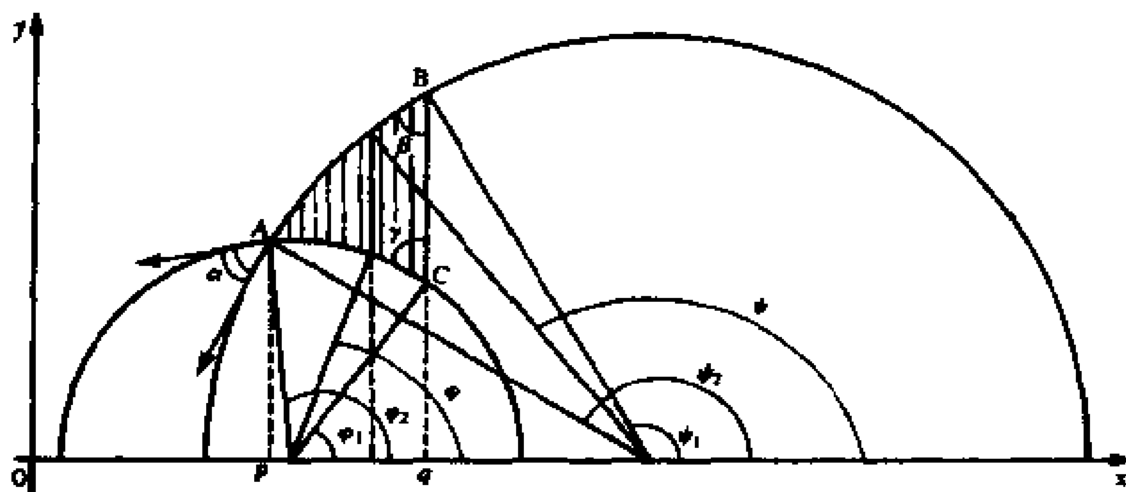


图 57

在圆弧$\overset{\frown}{CA}$上有$x=r\cos\varphi$,而在圆弧$\overset{\frown}{BA}$上有$x=R\cos\psi$.因此

$$\int_p^q \frac{\mathrm{d}x}{r\sin\varphi}=\varphi_2-\varphi_1,\ \int_p^q \frac{\mathrm{d}x}{R\sin\psi}=\psi_2-\psi_1.$$

这样

$$\begin{aligned} ABC\ \text{的面积} &= \varphi_2-\varphi_1-(\psi_2-\psi_1) \\ &= (\varphi_2-\psi_2)-\varphi_1+\psi_1. \end{aligned}$$

但

$$\alpha=\psi_2-\varphi_2,\ \gamma=\varphi_1,\ \pi-\beta=\psi_1,$$

所以最后得到

$$ABC\ \text{的面积}=\pi-(\alpha+\beta+\gamma).$$

最后,我们来证明只有一条非欧线垂直于两"平行"非欧线.在欧几里得"词典"中,这两条"直线"是直径分别为AA',BB'且没有公共点的两个半圆(图58).

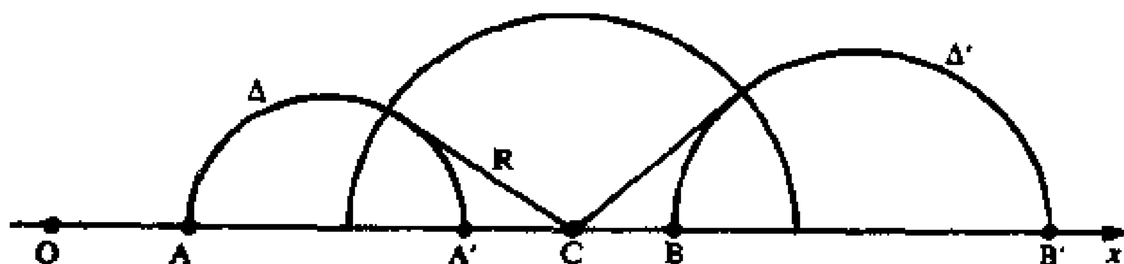


图 58

如果这两个半圆同心,则此唯一的"公垂线"是通过公共圆心平行于Oy的半直线.否则,作圆心在Ox上且与所给两半圆正交的半圆,设C是其圆心,R是其半径.如果x,a,a',b,b'分别是C,A,A',B,B'的横坐标,则必有

[255]

$$R^2=\overline{CA}\cdot\overline{CA'}=\overline{CB}\cdot\overline{CB'},$$

因之

$$R^2=(x-a)(x-a')=(x-b)(x-b'). \tag{13}$$

这说明 x 是一次方程

$$(a+a'-b-b')x=bb'-aa'$$

的解.由于所给两半圆不同心,系数 $a+a'-b-b'$ 不等于零.当把 x 的这个值代入(13)时,右边两表达式必为正数,因为否则就会有一正实数 y 满足

$$(x-a)(x-a')=(x-b)(x-b')=-y^2,$$

而这会蕴涵两个半圆

$$(x-a)(x-a')+y^2=0,(x-b)(x-b')+y^2=0$$

在半平面 H 中有一公共点,这与假定矛盾.

2. 实数模型

A. 有 理 数 理 论

实数公理(见第三章附录2)蕴涵在 $\mathbf{R}$ 内存在有理整数的加法子群 $\mathbf{Z}$ 和有理数 $p/q(p\in\mathbf{Z},q\in\mathbf{Z}$ 且 $q\neq0)$ 的子域 $\mathbf{Q}$.为在自然数理论中得到实数的模型,我们首先在自然数理论中构造 $\mathbf{Z}$ 和 $\mathbf{Q}$ 的模型.

对于 $\mathbf{Z}$,在 $\mathbf{N}^*\times\mathbf{N}^*$ 中考虑下述等价关系:两个偶 (p,q) 与 (p',q') 等价,如果

$$p+q'=p'+q;$$

我们以 $Z(\mathbf{N}^*)$ 记关于此关系的等价类的集合.

存在 $\mathbf{Z}$ 到 $Z(\mathbf{N}^*)$ 上的一个典范一一映射 d.为定义此映射,如果 $z>0$,令 $(z+1,1)$ 对应于 $z\in\mathbf{Z}$,如果 $z<0$,则令 $(1,1-z)$ 对应于 $z\in\mathbf{Z}$(我们记得我们已使 $\mathbf{N}^*$ 等同于实数 $n\cdot1(n\in\mathbf{N}^*)$ 的集合).定义 $d(z)$ 是所说偶的等价类.这就得到我们寻求的模型,其验证是简单的.

对于 $\mathbf{Q}$,通过类似的步骤,我们在整数理论中构造其模型(由此如果我们愿意,就可在自然数理论中构造其模型):考虑满足 $q\neq0$ 的整数偶 (p,q) 的集合 $C(\mathbf{Z})$ 以及下述等价关系:两个

偶(p,q)与(p',q')等价,如果

$$pq' = p'q,$$

令 $Q(\mathbf{Z})$为关于此等价关系的等价类的集合.通过令偶(p,q)的等价类 $c(p/q)$对应于有理数 $p/q(p,q\in\mathbf{Z},q\neq 0)$,可定义 $\mathbf{Q}$ 到 $Q(\mathbf{Z})$上的典范一一映射 c.由定义立即得知,如果 $p/q=p'/q'$,则 $c(p/q)=c(p'/q')$,因而我们定义了所需的模型.

256

B. 戴德金模型(简述)

这是 $\mathbf{R}$ 在有理数理论中的一个模型(由 A,它也是在自然数理论中的模型).在有理数集所有子集的集合 $\mathfrak{P}(\mathbf{Q})$中,定义如下的子集 $X\subset Q$ 构成的子集 $\mathfrak{R}(\mathbf{Q})$: $X\neq\varnothing$,$X\neq\mathbf{Q}$,且 X 具有下述性质: X 没有最大元,且如果 $x\in X$,$y<x$,则有 $y\in X$.

为简单起见,称 $\mathfrak{R}(\mathbf{Q})$的元素为 $\mathbf{Q}$ 的截段.

在 $\mathfrak{R}(\mathbf{Q})$中可给出由包含关系 $X\subset Y$ 定义的序结构.它在 $\mathfrak{R}(\mathbf{Q})$上是一个全序结构.事实上,假定 X,Y 是两个不同的截段,则或者存在 $x\in X$ 使得 $x\notin Y$,或者存在 $y\in Y$ 使得 $y\notin X$.在第一种情形,有理数 $u\geqslant x$ 不可能属于 Y,否则由定义就会有 $x\in Y$,从而每个 $y\in Y$ 满足条件 $y<x$,由定义这蕴涵 $y\in X$,因而 $Y\subset X$.同样的论证可得第二种情形导致 $X\subset Y$.

对每个实数 x,定义 S_x 为由有理数 $y<x$ 构成的截段.于是映射 $x\longmapsto S_x$ 是 $\mathbf{R}$ 到 $\mathfrak{R}(\mathbf{Q})$上的一一映射,它并且是有序集之间的同构.事实上,如果在 $\mathbf{R}$ 中 $x\neq y$,例如,$x<y$,则由阿基米德公理(第三章附录 2),存在有理数 r,使得 $x<r<y$,因而有 $S_x\subset S_y$ 且 $S_x\neq S_y$.另一方面,如果 $S\in\mathfrak{R}(\mathbf{Q})$是任一截段,则存在有理数 r,使得 $S\subset S_r$,因为否则就会有 $S=\mathbf{Q}$.作为区间套公理的推论,在 $\mathbf{R}$ 中存在 S 的最小上界 x(S 的上确界),因而我们

此时满足

$$a_{p(2m-2)} < x < a_{p(2m-1)}$$

的实数 x 构成一个无穷集,于是令 $p(2m)$为满足 $k > p(2m-1)$ 和

$$a_{p(2m-2)} < a_k < a_{p(2m-1)}$$

的最小整数 k.

4) 同样,$p(2m+1)$是满足 $k > p(2m)$与

$$a_{p(2m)} < a_k < a_{p(2m-1)}$$

的最小整数 k.

由定义显见对每个 $n \geqslant 0$ 有 $p(n+1) > p(n)$. 用归纳法可证对每个 $n \geqslant 0$ 有 $p(n) \geqslant n$. 这样我们就定义了闭区间

$$[a_{p(2m)}, a_{p(2m+1)}]$$

构成的序列,使得其中每一个都包含于前一个区间之中. 更精确地说,对每个 $m \geqslant 1$ 有

$$a_{p(2m-2)} < a_{p(2m)} < a_{p(2m+1)} < a_{p(2m-1)}. \tag{1}$$

作为区间套公理的推论,存在实数 y,它属于所有这些区间,而由(1),y 不可能与任一区间的端点相同.

由假定,存在自然数 q,使得 $y = a_q$. 由于序列$(p(n))$无限递增,所以存在最大的整数 n,使得 $p(n) \leqslant q$;因此又有 $q < p(n+1)$. 先假定 n 是偶数:$n = 2m$,于是我们就会有

$$a_{p(2m)} < a_q < a_{p(2m+1)} < a_{p(2m-1)},$$

从而 $p(2m+1)$就不是满足 $k > p(2m)$与 $a_{p(2m)} < a_k < a_{p(2m-1)}$ 的最小整数 k. 另一方面,假定 n 是奇数:$n = 2m-1$,我们就会有

$$a_{p(2m-2)} < a_{p(2m)} < a_q < a_{p(2m-1)},$$

[258] 从而 $p(2m)$就不是满足 $k > p(2m-1)$与 $a_{p(2m-2)} < a_k < a_{p(2m-1)}$的最小整数 k. 两种情形都导致矛盾.

B. 基数之间的序关系

设 E,F 是任意两个非空集，$f:E\to F$ 是 E 到子集 $f(E)\subset F$ 上的一一映射，$g:F\to E$ 是 F 到子集 $g(F)\subset E$ 上的一一映射. 问题在于证明 E 与 F 是等势的.

令 $h=g\circ f:E\to E$，因而 h 是 E 到一个子集 $h(E)\subset E$ 上的一一映射. 令 $R=E\setminus g(F)$并考虑这样的性质：对 $M\in\mathfrak{P}(E)$，有

$$M\supset R\cup h(M)$$

(右边表示 R 与 $h(M)$之并)；令 $\mathfrak{F}\subset\mathfrak{P}(E)$是具有上述性质的元素 $M\in\mathfrak{P}(E)$的集合. 集合 $\mathfrak{F}$ 非空，因为显然 $E\in\mathfrak{F}$. 设 A 是所有集合$M\in\mathfrak{F}$ 之交. 这个集合也属于 $\mathfrak{F}$. 事实上，我们有 $R\subset A$，因此对所有 $M\in\mathfrak{F}$ 有 $R\subset M$；再者对每个 $M\in\mathfrak{F}$ 有 $h(A)\subset h(M)$，因之对每个 $M\in\mathfrak{F}$ 有

$$R\cup h(A)\subset R\cup h(M)\subset M,$$

由定义，这蕴涵 $R\cup h(A)\subset A$.

其次，我们来证明 $A=R\cup h(A)$. 用反证法，假定存在 $x\in A$，但 x 不属于$R\cup h(A)$. 令 $M=A\setminus\{x\}$. 由于 $x\notin h(A)$，故有

$$h(A)\subset A\setminus\{x\}=M,$$

另一方面，$h(M)\subset h(A)$. 于是我们就会有 $R\cup h(M)\subset M$，这与 A 的定义矛盾.

令 $A'=f(A)$，$B'=F\setminus f(A)$，$B=g(B')$. 由 h 的定义，得 $g(A')=h(A)$. 由于 g 是 F 到$g(F)$上的一一映射且由于 $A'\cap B'=\varnothing$，所以 $B\cap g(A')=\varnothing$，这就是说，$B\cap h(A)=\varnothing$. 由于 $B\cap R=\varnothing$，所以 $B\cap A=\varnothing$.

我们来证明 B 等于A 的补集$E\setminus A$. 事实上，取 $x\in E\setminus A$. 由定义，存在 $M\in\mathfrak{F}$，使得 $x\notin M$，于是更有 $x\notin h(M)$. 另一方面，$x\notin R$，这表明 $x\in g(F)$. 设 $y\in F$ 是使得$x=g(y)$的唯一元素；我们来证明必有 $y\in B'$从而$x\in B$. 如其不然，就会有

$$y\in A'=f(A),$$

换言之,就会对某个元素 $u\in A$ 有 $y=f(u)$,因而

$$x=g(f(u))=h(u)\in h(A).$$

但因 $h(A)\subset h(M)$,我们就会有 $x\in h(M)$;然而我们已证明 $x\notin h(M)$(图 59).

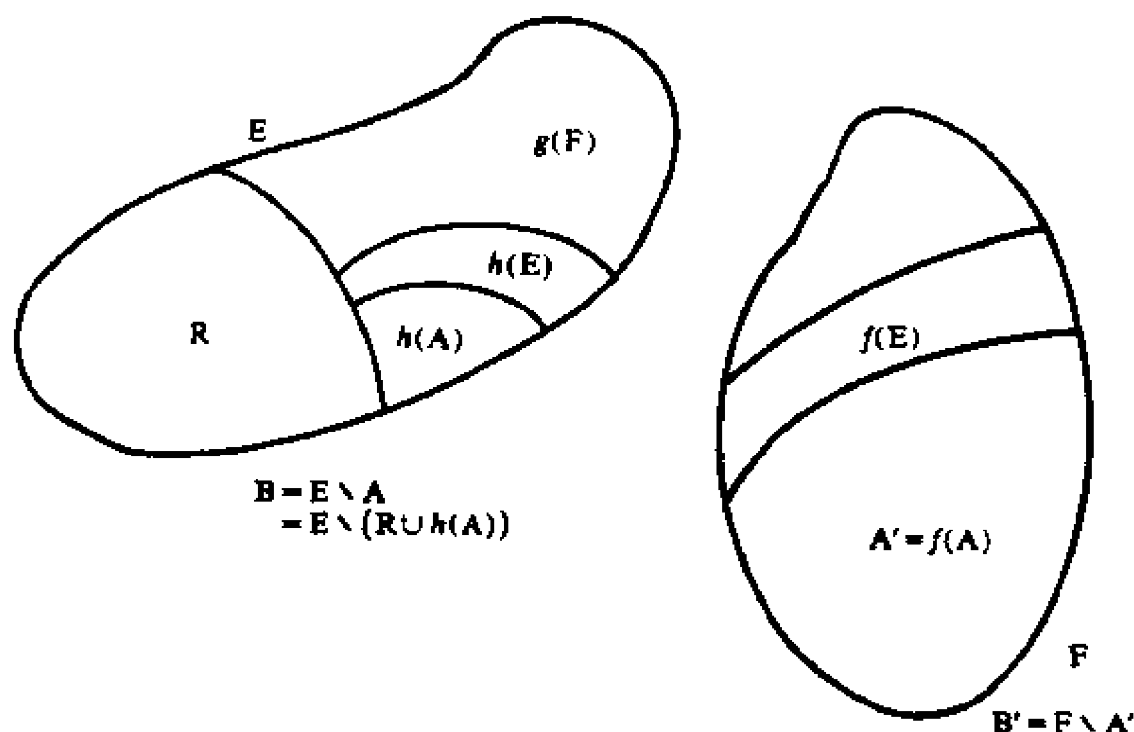


图 59

这样我们把 E 分解为两个没有公共元素的子集之并 $A\cup B$,又把 F 分解为两个没有公共元素的子集之并 $A'\cup B'$.此外,f 在 A 上的限制 f_1 是 A 到 A' 上的一一映射,g 在 B' 上的限制 g_1 是 B' 到 B 上的一一映射.对 $x\in A$ 取 $\Phi(x)=f_1(x)$,对 $x\in B$ 取 $\Phi(x)=g_1^{-1}(x)$,显然我们就定义了 E 到 F 上的一一映射 Φ.

[259] 注意上述证明并未用到无穷集的存在性.

C. R 与 $\mathbf{R}^2=\mathbf{R}\times\mathbf{R}$ 等势

证明分四步进行.

（Ⅰ）集合 $\mathbf{R}$ 与满足 $0<t<1$ 的数 t 构成的开区间 I 等势.为此只须注意

$$t \longmapsto \tan\pi\left(t-\frac{1}{2}\right)$$

是 I 到 $\mathbf{R}$ 上的一一映射.

Ⅱ）集合 I 与属于 I 的无理数的子集 $J=I\setminus(I\cap\mathbf{Q})$ 等势.我们应用上面 B 的结论.首先 $J\subset I$.另一方面,设 θ 是一无理数,满足 $1<\theta<2$.数 $t+\theta$(其中 $t\in I\cap\mathbf{Q}$)的集合 A 由无理数构成,且包含于开区间 $0<t<3$ 之内,因而也包含于属于该区间的无理数的集合 $3J$ 之内.但 $t\longmapsto 3t$ 是 J 到 $3J$ 上的一个一一映射;另一方面,I 是没有公共元素的集合 J 与 $I\cap\mathbf{Q}$ 之并,因此 I 与 $J\cup A\subset 3J$ 等势,由此得知 I 与 J 的一个子集等势.

Ⅲ）第三步是本证明的关键步骤,它是要证明 J 与自然数的所有序列(a_n)的集合 S(也记为$(\mathbf{N}^*)^{\mathbf{N}^*}$)等势.$J$ 到 S 上的一一映射通过连分数理论来定义.对每个无理数 $x\in J$,通过对 n 进行递推来对应自然数的一个序列(a_n).由于

$$\frac{1}{x}>1,$$

我们可写

$$\frac{1}{x}=a_1+x_1,$$

[260]

其中 $a_1\in\mathbf{N}^*$,$0<x_1<1$ 因之 $x_1\in J$.然后递推地进行,置

$$\frac{1}{x_1}=a_2+x_2,$$

$$\cdots\cdots$$

$$\frac{1}{x_{n-1}}=a_n+x_n,$$

$$\cdots\cdots$$

其中对每个 n,a_n 是自然数,且 $0<x_n<1$.递推过程不会在有限步终止,因为任何 x_n 都非有理数:如果 x_n 是有理数,那么

$x_{n-1}, x_{n-2}, \cdots, x_1$ 就会是有理数,最后 x 就会是有理数,与假设矛盾.这样,我们对每个 $x \in J$ 定义了自然数的一个无穷序列 $s(x) = (a_n)_{n \geq 1}$.

我们还必须证明,对每个序列 $(a_n)_{n \geq 1}$,存在唯一的 $x \in J$,使得 $s(x) = (a_n)_{n \geq 1}$.这是欧拉通过下面的论证所证明的.首先对每个 n 考虑部分有限序列 $(a_1, a_2, \cdots, a_n)$,并由下列条件定义属于 I 的有理数的一个有限序列:

$$
\begin{aligned}
\frac{1}{y_1} &= a_n, \\
\frac{1}{y_2} &= a_{n-1} + y_1, \\
&\cdots\cdots \\
\frac{1}{y_{n-1}} &= a_2 + y_{n-2}, \\
\frac{1}{y_n} &= a_1 + y_{n-1}.
\end{aligned}
\tag{2}
$$

(注意,如果 $y_n = u/v$,其中 u, v 是自然数且 $u < v$,则把上面的方程从下读到上,它们正是计算 u 与 v 的最大公因数的欧几里得除法的逐次序列的另一写法(参见第四章附录3).事实上,

$$
\begin{aligned}
v &= a_1 u + r_1, & r_1 &< u, \\
u &= a_2 r_1 + r_2, & r_2 &< r_1, \\
r_1 &= a_3 r_2 + r_3, & r_3 &< r_2, \\
&\cdots\cdots \\
r_{n-3} &= a_{n-1} r_{n-2} + r_{n-1}, & r_{n-1} &< r_{n-2} \\
r_{n-2} &= a_n r_{n-1}, & r_{n-1} &\equiv \{u, v\},
\end{aligned}
$$

其中 $\{u, v\}$ 是 u 与 v 的最大公因数.自然数 a_j 即为出现于(2)中的 a_j,而 $y_j (1 \leq j \leq n-1)$ 由

$$
y_{n-j} = \frac{r_j}{r_{j-1}}
$$

给定.)

数 y_n 记作

$$\frac{1|}{|a_1}+\frac{1|}{|a_2}+\cdots+\frac{1|}{|a_n}, \tag{3}$$ [261]

并称 $a_1,a_2,\cdots,a_n$ 为 y_n 的部分商，而(3)称为 y_n 的连分数展开.可以证明，如果自然数 $p_j,q_j(1\leqslant j\leqslant n)$ 由下述公式递推地定义：

$$\begin{aligned}&p_1=1,q_1=a_1,p_2=a_2,q_2=1+a_2a_1,\\&p_j=p_{j-2}+a_jp_{j-1},q_j=q_{j-2}+a_jq_{j-1},j\geqslant 3,\end{aligned}\tag{4}$$

则有 $y_n=p_n/q_n$.如果我们回到无穷序列 $(a_n)_{n\geqslant 1}$，则可递推地对每个自然数 j 定义 p_j,q_j.欧拉用归纳法证明，对每个 $n\geqslant 2$，有

$$\frac{p_n}{q_n}-\frac{p_{n-1}}{q_{n-1}}=\frac{(-1)^{n-1}}{q_nq_{n-1}}, \tag{5}$$

由此从对所有 n 有 $a_n\geqslant 1$ 这个事实出发，从(4)用归纳法可证对所有 n 有 $q_n\geqslant n$.于是

$$\left|\frac{p_n}{q_n}-\frac{p_{n-1}}{q_{n-1}}\right|\leqslant\frac{1}{n(n-1)}.$$

由于通项为 $1/n(n-1)$ 的级数收敛，所以序列 (p_n/q_n) 趋于一个极限 x.再者，由(5)式，对于 $m\geqslant 1$，闭区间

$$\left[\frac{p_{2m}}{q_{2m}},\frac{p_{2m+1}}{q_{2m+1}}\right]$$

是包含于 I 内的区间套，从而 $x\in I$.我们以表示式

$$\frac{1|}{|a_1}+\frac{1|}{|a_2}+\cdots+\frac{1|}{|a_n}+\cdots \tag{6}$$

记数 x，称为 x 的(无限)连分数展开，而有理数 p_n/q_n 称为此展开的渐近分数.我们还看到，如果

$$x_1=\frac{1|}{|a_2}+\frac{1|}{|a_3}+\cdots+\frac{1|}{|a_n}+\cdots,$$

则有

$$\frac{1}{x}=a_1+x_1.$$

由递推可知 x 是无理数,而 x 的部分商序列 $(a_n)_{n\geq 1}$ 确是序列 $s(x)$.

Ⅳ) 我们现在就可来证明集合 $S\times S$ 与集合 S 等势. 为此,康托尔简单地观察到,映射

$$((b_n),(c_n))\longmapsto(a_n)$$

(其中对所有 $k\geq 1$ 令 $a_{2k-1}=b_k, a_{2k}=c_k$)是 $S\times S$ 到 S 上的一个一一映射.

D. 子集的集合的基数

[262] 设 E 是任一集合,映射 $x\longmapsto\{x\}$ 使 E 的子集 $\{x\}$——$\mathfrak{P}(E)$ 的一个元素,其唯一成员为 x[①]——对应于每个元素 $x\in E$,它是 E 到 $\mathfrak{P}(E)$ 的一个子集上的一一映射. 我们还必须证明不存在 E 到 $\mathfrak{P}(E)$ 上的一一映射. 用反证法,假定存在这样的映射 $f: E\to\mathfrak{P}(E)$. 考虑这样的性质:对元素 $x\in E$,有 $x\notin f(x)$. 由策梅鲁概括公理(本章 §4,C),存在由具有此性质的 $x\in E$ 所构成的子集 $X\subset E$. 我们来证明 $X\notin f(E)$,而这与 $f(E)=\mathfrak{P}(E)$ 这个假定相矛盾. 如果 $x\in X$,由定义,$x\notin f(x)$,因此 $f(x)\neq X$. 另一方面,如果 $x\notin X$,则由定义有 $x\in f(x)$,因此又有 $f(x)\neq X$. 这样 X 不可能是 $f(E)$ 的元素.

注意我们在上述证明中既没有用无穷公理,也没有用选择公理.

当 E 是具有 n 个元素的有限集时,通过对 n 进行归纳可证 $\mathfrak{P}(E)$ 具有 2^n 个元素,这解释了一般记号 $2^{\mathfrak{m}}$($\mathfrak{m}$ 是任一个基数)[263] 的缘起.

---

① 把一个对象 x 与由单个元素 x 构成的集合 $\{x\}$ 加以区别(此点应归于弗雷格)看来或许有点故弄玄虚,但却被证明在集合论的论证中是有用的,这里只是一个非常简单的例子. 在更复杂的证明中,为避免混淆,始终知道哪个集合含有所研究的对象是重要的.

附 录
数学家小传

读者在本小传中可以得到关于本书引用过的数学家的某些信息. 原则上我们限于1985年前逝世的数学家. 如果想得到这方面完备和内容充实的资料, 读者可参阅《科学家传记辞典》(Dictionary of Scientific Biography, 16卷, New York(Ch. Scribner's Sons), 1970—1980)[①].

阿贝尔*(Abel, Niels Henrick; 1802—1829)

阿贝尔生于挪威芬岛, 就读于奥斯陆大学. 1825年得政府资助, 游学柏林和巴黎. 因患肺结核卒于挪威弗鲁兰. 他独立于高斯和雅可比, 是椭圆函数论的开创者之一, 并建立了阿贝尔积分理论. 归于他名下的还有分析中的许多定理以及关于代数方程(现称为"阿贝尔方程")可用根式求解的发现.

达朗贝尔*(D'Alembert, Jean le Rond; 1717—1783)

达朗贝尔生于法国巴黎, 是一对贵族男女的私生子. 在四国学院学习法律、医学和数学. 1741年当选为巴黎科学院院员[②].

---

① 也可参阅吴文俊主编的《世界著名数学家传记》, 上下册, 科学出版社, 1995年. 以下开列的数学家中, 凡出现于此书者, 译者均附加 * 号. 读者可从此书读到关于该数学家生平、工作、著述和思想的全面准确的记述. 又, 小传中大多数地名均未说明国别, 译者均已补上. ——译注

② 当时巴黎科学院有四种职称: 荣誉院士、终身院士、副院士、助理院士. 严格讲只有前两种才是正式院士. 达朗贝尔开始时是助理院士. ——译注

1772 年成为法兰西学院终身秘书.17 世纪 50 年代,是狄得罗主编的《百科全书》的科学编辑.他在分析和力学中留下了大量论著.

亚历山大(Alexander, James Waddell; 1888—1971)

亚历山大生于美国新泽西州海辉市.1928 至 1933 年执教于普林斯顿大学,1933 年后执教于普林斯顿高等研究院.他是代数拓扑领域中一些重要著作的作者.

花拉子米*(al - Khwārizmī, Abū Ja'far Muhammad Ibn Mūsā; 800 年前生,847 年后卒)

花拉子米可能生于伊朗,是阿拔斯哈利发创办的科学机关[1]的成员;撰写了天文学和代数学著作,后者虽无独创内容,[267] 但在中世纪有非常大的影响.

安培(Ampere, André - Marie; 1775—1836)

安培生于法国里昂,以在里昂教授数学开始其科学生涯.1802 年为布尔格中心学校教师,1803 年成为综合工科学校辅导教师,1824 年为法兰西学院实验物理教授.自 1814 年起为法兰西研究院成员.他以其物理学方面的工作著称于世.

阿波罗尼奥斯*(Apollonius; 公元前 3 世纪中至公元前 2 世纪初)

据称阿波罗尼奥斯生于小亚细亚的一个希腊小村佩尔格,对其生平所知甚少.他定居于亚历山大,并常访问帕加马和以弗所.撰写了关于圆锥曲线的极其完备的专著.

---

① 即著名的"智慧馆"(Bayt al - Hikmah),是继公元前 3 世纪亚历山大博物馆之后最著名的学术机关.——译注

阿基米德*(Archimedes;公元前287—212)

阿基米德定居于叙拉古,极有可能访问过亚历山大,并在那里与欧几里得的一些继承者共同工作.但他必定返回叙拉古,在那里负责港口设备与海军和军事建筑工程,并在叙拉古为罗马军队攻陷时被杀.他的著作是古典时代最重要的文献,他还是微积分的直接先驱者.

阿尔冈(Argand,Jean Robert;1768—1822)

阿尔冈生于瑞士日内瓦,作为一名会计在法国巴黎度其一生.对其生平我们几乎一无所知.他自学数学.

巴拿赫*(Banach,Stefan;1892—1945)

巴拿赫生于波兰的克拉科夫,就读于里沃夫(属乌克兰)工学院,于1927年成为该院教授.他是泛函分析奠基者之一,证明了这一领域中的一些最重要的定理.

贝尔特拉米*(Beltrami,Eugenio;1835—1899)

贝尔特拉米生于意大利克雷莫纳,在帕维亚和米兰接受大学教育.为波伦亚大学(1862—1864,1866—1873)、比萨大学(1864—1866)、罗马大学(1873—1876,1891—1899)和帕维亚大学(1876—1891)教授.1898年任林琴学院院长.其最重要的著作为论述微分几何的论著.

丹尼尔·伯努利*(Bernoulli,Daniel;1700—1782)

丹尼尔·伯努利生于荷兰格罗宁根,其父约翰·伯努利为当地数学教授.从1705年起居住于瑞士巴塞尔,其间于威尼斯旅居(1723—1724),又在圣彼得堡就过8年(1725—1733)并在彼得堡科学院工作.1733至1776年执教于巴塞尔大学,初为植物学和解剖学教授,后为物理学教授.其最重要的著作为论述流体

[268] 动力学、概率论和弦振动方程的论著.

雅格布·伯努利*(Bernoulli,Jacob;1654—1705)

雅格布·伯努利生于瑞士巴塞尔,在巴塞尔大学攻读哲学、神学、数学和天文学,研习后两项违背其父意愿.在游学法国、荷兰、英国后,从1683年起执教于巴塞尔大学,特别于1687年后为该校数学教授.作为莱布尼茨的学生,他开发了微积分学在级数论、变分法、概率论以及力学中的众多应用.

约翰·伯努利*(Bernoulli,John;1667—1748)

约翰·伯努利是雅格布·伯努利之弟,生于瑞士巴塞尔,打算攻读医学,但实际上随其兄研习数学.1691—1692年居于巴黎,1695年任荷兰格罗宁根大学数学教授,1705年继任其兄在巴塞尔大学的职务.他比其兄较少独创性,但同其兄一样,为微积分学及其应用最积极的传播者之一.

伯恩斯坦(Bernstein,Felix;1878—1956)

伯恩斯坦生于德国哈雷,在该城为康托尔的学生,然后在格丁根大学师从希尔伯特和克莱因,并于1911年任该校统计数学教授,直到1934年纳粹政权撤销他的职务为止.以后他在美国的几所大学执教.他的工作领域是集合论和数学遗传学.

贝祖(Bezout,Etienne;1739—1783)

贝祖生于法国内穆尔.1758年起为巴黎科学院成员.1763年在近卫军和海军以及炮兵部队任未来军官的数学教官和考试官.其著作涉及代数方程及其对早期代数几何的应用.

波尔约*(Bolyai,János;1802—1860)

波尔约生于匈牙利科罗日瓦[1],其父为数学教师并为高斯的同窗,但波尔约从事军事工作.1833年因伤残退伍,先回其父处,然后迁居多马尔德.当他得知高斯在他之前已发现他所独立发现的非欧几何后,就再也没有发表任何数学论著.

波尔查诺(Bolzano,Bernard;1781—1848)

波尔查诺生于捷克布拉格,并在布拉格大学攻读神学和数学.1804年被委任为神父,1805年任奥地利皇帝刚开设的布拉格大学哲学和宗教教授.1819年因异端思想被撤职,其著作被列入天主教禁书目录.他的独创思想远远领先于他所处的时代, [269] 构思了许多后来才得到发展的观念,但由于缺少专业工具而未能自己提炼出这些观念的所有内涵.

邦贝利(Bombelli,Rafael;1526—1572)

邦贝利生于意大利波伦亚,职业为工程师.他关于代数的专著比卡尔达诺的有关著作更加清晰系统,以其所用记号以及对于"虚数"的讨论而影响后代数学家.

布尔*(Boole,George;1815—1864)

布尔生于英国林肯市,自学数学,16岁时开始当教师,1835年在林肯市创办一所中学.1849年起任才建立的科克女皇学院(爱尔兰)教授.1857年被选为英国皇家学会会员.他应看作数理逻辑的奠基者.

波莱尔*(Borel,Emile;1871—1956)

波莱尔生于法国阿韦龙省圣·阿弗里克,就读于高等师范学校.执教于里尔大学和高等师范学校,1909年起执教于巴黎大

① 今罗马尼亚克卢日.——译注

学.1924 至 1936 年为国民议会议员,1925 年任海军部长.他创建国家科学研究中心,协助筹建亨利·庞加莱研究所,并从 1928 年直止逝世一直担任该所所长.1921 年入选巴黎科学院.他是最早把康托尔的观念应用于分析和概率论的数学家之一.

波罗摩笈多 *(Brahmagupta;598 至 665 后)

波罗摩笈多可能居住于印度拉贾斯坦邦.有关于天文学和数学的若干著作留传于世.

布劳威尔 *(Brouwer,Luitzen Egbertus Jan;1881—1966)

布劳威尔生于荷兰奥弗希,毕生工作于阿姆斯特丹大学,1907 年在该校获博士学位,1909—1951 年执教于该校.他于 1910—1912 年发表的关于代数拓扑的论著使他扬名天下,但嗣后终其一生致力于发展"直觉主义者"的数学.

布尼亚科夫斯基(Buniakowski,Victor Jakovlevich;1804—1889)

布尼亚科夫斯基于 1825 年在巴黎获博士学位,自 1846 年起执教于圣彼得堡大学.1828 年入选彼得堡科学院.

坎帕纳斯(Campanus;13 世纪前 25 年中至 1296)

坎帕纳斯可能生于意大利诺瓦拉,是关于天文学的古代著[270] 作和欧几里得《几何原本》的编纂者和评注者.

康托尔 *(Cantor,Georg;1845—1918)

康托尔生于俄国圣彼得堡,双亲为德国人.起初就读于苏黎世大学,后入柏林大学,师事魏尔斯特拉斯.1869 年起执教于哈雷大学.1890 年创建德国数学家联合会,并任首任主席.1897 年筹备在苏黎世召开的第一届国际数学家大会.1884 年后不时发

作深度忧郁症，最后死于哈雷大学精神病诊所．他以其关于集合论和拓扑学的论著使自己成为其思想对当代数学发生最大影响的人物之一．

卡尔达诺*（Cardano，Girolamo；1502—1576）

卡尔达诺生于意大利帕维亚，在帕维亚和帕多瓦两地攻读医学．1534年起在米兰的一所学校任数学教师，同时行医．1543年任帕维亚大学医学教授，以后又于1562年任波伦亚大学医学教授．1570年突然被判犯异端邪说罪，被逮捕并被撤销职务．他是百科全书式的学者，其对代数方程理论的贡献首次为这一理论带来一般观念．

埃利·嘉当*（Cartan，Elie；1869—1951）

嘉当生于法境阿尔卑斯山的多洛米约．获政府助学金得以就读于里昂公立中学，后就读于高等师范学校．执教于蒙彼利埃大学、里昂大学、南锡大学和巴黎大学．1931年当选为巴黎科学院院士．他以其非凡独创的观念和技巧，对现今称作流形上的分析这一领域及其在李群论、整体微分几何和微分拓扑、偏微分方程以及力学中的众多应用作出了最大贡献．

柯西*（Cauchy，Augustin - Louis；1789—1857）

柯西生于法国巴黎，在考入综合工科学校后，还就读于道桥学校，并被委任为市政工程的工程师．1813年返回巴黎，1815年起执教于综合工科学校，稍后执教于巴黎理学院和法兰西学院．1816年被任命为巴黎科学院院士．1830年革命后流亡都灵，后至布拉格，任查理十世之孙的宫廷教师．1838年返回法国重获在科学院中的职务，1848年重获巴黎大学教授职位．柯西是欧拉之后最多产的数学家，像欧拉一样在纯粹数学和应用数学的所有领域多有建树．他是19世纪前半期分析的法典制订者，对

[271] 其进展贡献最大.他还是弹性论的奠基人.

谢瓦莱(Chevalley,Claude;1909—1984)

谢瓦莱生于约翰内斯堡(南非德兰士瓦省),是时其父为法国总领事.1926—1929年就读于高等师范学校.1936—1938年执教于斯特拉斯堡大学和雷恩大学;后赴美国,执教于普林斯顿大学(1940—1948)和纽约哥伦比亚大学(1948—1955).返法后自1955年起任巴黎大学教授直至1978年退休.他是法国科学院通讯院士,对代数数论特别是代数群理论作出了重大贡献.

许凯(Chuquet,Nicolas;15世纪下半期)

许凯是《数的科学中的三个部分》的作者,此书以其新颖的符号体系而引人注目.对其生平几乎一无所知.

克莱罗(Clairaut,Alexis - Claude;1713—1765)

克莱罗生于法国巴黎,其父为数学教授.他异常早熟,18岁时即入选巴黎科学院.他开创了关于挠曲线的研究,对18世纪中微积分的发展及其在天体力学中的应用颇多贡献.

科茨(Cotes,Roger;1682—1716)

科茨生于英国布巴奇,就读于剑桥大学,于1711年被选为英国皇家学会会员.他曾协助牛顿出版《自然哲学的数学原理》第二版,并写有关于"初等"函数积分的有趣论文,其中首次出现表示 $\cos\varphi + i\sin\varphi$ 的对数的公式.

戴德金*(Dedekind,Julius Wilhelm Richard;1831—1916)

戴德金生于德国不伦瑞克,就读于格丁根大学.1854—1855年开始作为无薪讲师在格丁根任教并与狄利克雷和黎曼密切合作.1858年去苏黎世综合工业大学,1862年任不伦瑞克高等工

业学院教授，直至去世．他与库默尔、克罗内克同为代数数论创立者，并与 H·韦伯合作首次提供了代数曲线的完整代数论述．与康托尔、希尔伯特同为数学中现代概念的开拓者．

德·费罗(Del Ferro,Scipione;1465—1526)

德·费罗生于意大利波伦亚，卒于同地，1496 至 1526 年间为波伦亚大学教授．未留论著(包括手稿)于世．

笛卡儿*(Descartes,René du Perron;1596—1650)

笛卡儿生于法国图赖纳省拉艾镇，曾获普瓦捷大学法律学位，并于巴黎在米多热和梅森指导下研习数学．1617 年参加奥伦治王子的军队，在嗣后 9 年中不时在各种军队中服役或在巴黎消耗时光．1628 年移居荷兰，1649 年应女皇克里斯蒂娜之邀 [272] 去瑞典，抵瑞典后不久即因患肺炎去世．他具有广博的才智，是各个领域的革新者；引进当今通用的代数符号系统后，独立于费马创立数学中的坐标方法．

丢番图*(Diophantus)

实际上对丢番图生平一无所知．他可能于公元 3 世纪中期活跃于亚历山大．

狄里克雷*(Dirichlet,Gustav Peter Lejeune;1805—1859)

狄里克雷生于德国迪伦，1822 至 1826 年在巴黎攻读，并以在法伊将军处任家庭教师维持生计．1826 至 1828 年在布雷斯劳大学任教，然后从 1829 至 1855 年在柏林大学任教．1855 年继任高斯在格丁根大学的职位．1831 年当选为柏林科学院院士．他在分析和代数数论中有重要论著，并开创了解析数论．

爱拉托斯特尼(Eratosthenes;公元前约 276—约 195)

爱拉托斯特尼生于施勒尼,在亚历山大度过其大部分生涯,是著名的亚历山大博物馆馆长.

欧几里得 * (Euclid;约公元前295年前后)

关于欧几里得生平所知极少.他可能在雅典求学并建立了亚历山大数学学派.

欧多克索斯 * (Eudoxus;公元前约400—约347)

欧多克索斯生于尼多斯,曾师事塔伦图姆的阿尔希塔斯和柏拉图,并旅游至埃及.他曾在基齐库斯教授学生,最后返回尼多斯并死于该地.他是古代世界最卓越的创新人物之一,几乎可以肯定是比例论和穷竭法的发明者.他创立了借助转动球面描述行星运行的第一个模型.

欧拉 * (Euler,Leonhard;1707—1783)

欧拉生于瑞士巴塞尔,1723年获巴塞尔大学博士学位,1726年接受圣彼得堡科学院的邀请,1727年移居该市,先为圣彼得堡科学院物理学教授,后为数学教授.1738年右眼失明.1741年离圣彼得堡去柏林,在嗣后25年中为柏林科学院数学部主任.与普鲁士国王的矛盾促使欧拉于1766年返回圣彼得堡,稍后双目失明.他是科学史中最多产的作家,其著作涉及所有科学和技术主题.他与拉格朗日一起,以其丰富多彩的发现主[273]导了18世纪的数学.

法格纳诺(Fagnano,Giulio Carlo del Toschi,1682—1766)

法格纳诺生于意大利锡尼加里亚一贵族家庭,1723年其家被提名为该市旌旗悬挂地.他自学数学,以其关于"双纽线"的结果引起欧拉注意,后者推广了法格纳诺的结果,发现了椭圆函数论中的第一批结论.

费马*(Fermat,Pierre de;1601—1665)

费马生于法国博蒙;其父为皮革商,富有家产,可供其子在图卢兹大学攻读法律.费马于 1631 年获奥尔良大学学士学位后,获得图卢兹议会顾问之职.1648 年后任卡斯特尔法庭法官.他无疑是 17 世纪最深刻的数学家,与帕斯卡共创概率论,在笛卡儿之前发现坐标方法.他最早提供确定平面曲线切线的一般方法;在数论中尤其显示出他的天才.

费拉里(Ferrari,Ludovico;1522—1565)

费拉里生于意大利波伦亚,是卡尔达诺的学生和合作者,执教于米兰和波伦亚.他发现了通过把四次方程归结为三次方程来用根式求解四次方程的方法.

傅里叶*(Fourier,Jean Baptiste Joseph;1768—1830)

傅里叶 9 岁时成为孤儿,1789 年在其家乡奥塞尔任教.1794 年被捕,罗伯斯庇尔被处决后获释,随即回巴黎就读于师范学校,该校建于是年,且于同一年关闭.1795 年他任综合工科学校助教,1798 年随蒙日在拿破仑埃及远征军中工作.1801 年回国后,拿破仑任命他为伊泽尔地区的行政长官.百日皇朝后,由于朋友相助,得以任塞纳统计局主任.1817 年成为巴黎科学院成员,1822 年任该院终身秘书.1827 年被选为法兰西学院院士.他的主要工作是创建热的理论,这导致他创建现称为傅里叶级数和傅里叶积分的理论.

弗伦克尔(Fraenkel,Adolf Abraham;1891—1965)

弗伦克尔生于德国慕尼黑,就读于慕尼黑大学、马尔堡大学、柏林大学和布雷斯劳大学.1916 至 1925 年执教于马尔堡大学,1928 至 1929 年执教于基尔大学,1929 至 1959 年执教于耶

[274] 路撒冷希伯来大学.其工作领域为逻辑和集合论.

弗雷歇* (Frechét, Maurice; 1878—1973)

弗雷歇生于法国马利尼,就读于巴黎高等师范学校.相继执教于普瓦捷大学(1910—1919),斯特拉斯堡大学(1920—1927)和巴黎大学(1928—1940);为巴黎科学院院士,在泛函分析中引进了许多新思想.

弗雷格* (Frege, Friedrich Ludwig Gottlob; 1848—1925)

弗雷格生于德国维斯马,就读于耶拿大学和格丁根大学,并于1873年获格丁根大学哲学博士学位.1879至1917年任耶拿大学哲学系教授,其工作领域为数理逻辑及其应用.

伽利略 (Galilei, Galileo; 1564—1642)

伽利略生于意大利比萨,在比萨大学攻读医学,但在校外自学数学.1589至1592年任比萨大学教授,1592至1610年任帕多瓦大学教授,后移居佛罗伦萨及其邻近地区.1633年遭宗教审判,嗣后被软禁.他是现代物理学之父,以实验为手段,以数学模型为指导研究物理学.

伽罗瓦* (Galois, Evariste; 1811—1832)

伽罗瓦生于法国布尔格·拉赖因,就读于巴黎路易·勒格兰皇家中学.1829年试图进综合工科学校学习,但未通过入学考试,最后被高等师范学校录取.早于1830年即因共和派观点而被校方开除,然后投身于政治斗争,终于被捕入狱.因决斗致死,详情不明.

高斯* (Gauss, Carl Friedrich; 1777—1855)

高斯生于德国不伦瑞克一贫穷家庭,1792年获不伦瑞克公

爵资助得以就读于不伦瑞克卡洛琳学院和格丁根大学(1795—1798).1799年获海尔姆斯台特大学博士学位;1807年任格丁根天文台台长,直至去世.1818至1825年负责汉诺威地理测量工作.他如同欧拉那样广博,其思想甚至比欧拉更加深刻,为数学的所有分支带来了新的生机,可惜未能发表其全部发现.

吉拉尔(Girard,Albert;1595—1632)

吉拉尔生于法国洛林的圣·米伊埃,但由于是新教徒,不得不迁居荷兰.可能就读于莱顿大学,后在奥伦治王子拿骚伯爵腓特烈·亨利的军队中任工程师.

[275]

哥德尔*(Gödel,Kurt;1906—1978)

哥德尔生于奥匈帝国的布尔诺(今属捷克),就读于维也纳大学.其博士论文于1929年被通过,1933年任无薪讲师.1939年离维也纳到美国普林斯顿高等研究院工作,1953年成为该院教授.

格拉斯曼*(Grassmann,Hermann;1809—1877)

格拉斯曼生于德国斯德丁[①],在柏林大学攻读神学,后在柏林的中学任教,1842年后在斯德丁大学预科任教.他关于线性代数和外代数的思想在生前未获理解;生命后期致力于研究语言学,成为备受尊重的梵文专家.

阿达玛*(Hadamard,Jacques;1865—1963)

阿达玛生于法国凡尔赛,就读于高等师范学校.毕业后执教于巴黎布丰公立中学(1890—1893)、波尔多大学(1893—1897)、巴黎大学(1897—1909)、法兰西学院(1909—1937)、综合工科学

① 今波兰什切青.——译注

校(1912—1937)以及中央工艺和制造学校(1920—1937).1912年成为巴黎科学院院士.其工作主要集中在分析及其应用上.

哈密顿*(Hamilton, William Rowan; 1805—1865)

哈密顿生于爱尔兰都柏林,于1823年进入都柏林三一学院.1827年,并无学位的哈密顿被任命为敦辛克天文台的皇家天文研究员.1832年成为爱尔兰皇家科学院院士,1837至1845年间为该院院长.在力学和光学中有重要论著.

哈代*(Hardy, Godfrey Harold; 1877—1947)

哈代生于英国克兰利,1896年入剑桥大学三一学院;1919年前一直在该院攻读和任教,1919年任牛津大学教授,1928—1929年访问普林斯顿,1931年重返剑桥任教授,直至1942年退休.他在分析和解析数论中发表(部分与李特尔伍德、拉马努金合作)了重要论著.

亨泽尔(Hensel, Kurt; 1861—1941)

亨泽尔生于德国柯尼斯堡,就读于波恩和柏林,1884年在柏林大学获博士学位.初在柏林大学任教,1901年后任马尔堡大学教授.他在数论中引进了富有独创性的思想,其重要性直到[276] 1920年后才为世人所认识.

埃尔米特*(Hermite, Charles; 1822—1901)

埃尔米特生于法国洛林地区的迪约兹,1842年考取综合工科学校,一年后因右腿残疾被除名,随后埃尔米特决心从事教师职业.1848至1876年执教于综合工科学校,1869至1897年为巴黎理学院教授.1856年当选为巴黎科学院院士.他在代数、数论和分析中都有同等卓越的工作.

希尔伯特*(Hilbert, David; 1862—1943)

希尔伯特生于德国柯尼斯堡,1880至1884年在柯尼斯堡大学学习(其间第二学期就读于海德堡大学).曾赴莱比锡和巴黎游学,1886年在柯尼斯堡大学取得无薪讲师资格.1895年任格丁根大学教授,直至退休.作为非常广博的领头数学家,他与庞加莱一起成为对20世纪数学最有影响的人物.

希庇阿斯(埃利斯的)(Hippias of Elis;约公元前400年)

对希庇阿斯的生平所知很少,只确知其著作的标题.他似乎是一位具有百科全书知识的"诡辩学者".有些评论家把发明称为"割圆曲线"的超越曲线归功于他,此曲线可用来进行三等分任意角和化圆为方的几何作图.

希波克拉底(希俄斯的)(Hippocrates of Chios;公元前5世纪后半期)

关于希波克拉底只限于从轶闻中得知一些情况.据信他写作了关于几何学的第一本专著.

赫尔德(Hölder, Otto Ludwig; 1859—1937)

赫尔德生于德国斯图加特,曾在柏林大学师从魏尔斯特拉斯,1882年在蒂宾根大学获博士学位.他先为格丁根大学无薪讲师,后任柯尼斯堡大学(1894—1899)和莱比锡大学(1899年后)教授.有关于分析和群论的有趣论著传世.

雅可比*(Jacobi, Carl Gustav Jacob; 1804—1851)

雅可比生于德国波茨坦,其父为犹太银行家.在其博士论文为柏林大学通过之后,为能获得就职资格而改宗基督教.1826年应邀赴柯尼斯堡大学.1843年因健康状况恶化不得不去意大利休养.1844年回国后在柏林大学任教.雅可比是柏林科学院

院士.他关于椭圆函数和超椭圆函数的发现最受人们称道,但也发表了许多关于分析、变分法和数论的论文.

杰文斯(Jevons,William Stanley;1835—1882)

杰文斯生于英国利物浦,就读于伦敦的大学学院.先执教于曼彻斯特和利物浦,然后自 1876 年起任教于大学学院.其论著涉及逻辑学和政治经济学.1872 年成为英国皇家学会会员.

[277]

若尔当*(Jordan,Camille;1838—1921)

若尔当生于法国里昂,就读于综合工科学校,在 1885 年之前职务为工程师.从 1873 年起直至 1912 年退休,同时在综合工科学校和法兰西学院任教.1883 年当选为巴黎科学院院士.他关于群论的论著在 19 世纪最后阶段主导了这一数学分支,而他的《分析教程》一直是几代分析教学的楷模.

开普勒(Kepler,Johannes;1571—1630)

开普勒生于符腾堡公国的魏尔城,就读于蒂宾根大学,然后在奥地利格拉茨讲授数学,但由于信仰新教被勒令离开格拉茨.1601 年在布拉格继第谷任鲁道夫二世的御前天文学家.马蒂阿茨皇帝(1612—1619)任命他为上奥地利数学家并令其定居林茨,此职位在费迪南二世时仍为开普勒所有.尽管有显赫的官方庇护人,开普勒的经济状况仍无保障,晚年过度劳累,贫病交迫,死于雷根斯堡.

基灵(Killing,Wilhelm Karl Joseph;1847—1923)

基灵生于德国布班克,在明斯特和柏林接受大学教育.任教于布里隆预科学校(1878—1882)、不伦茨堡预科学校(1882—1892)和明斯特大学(1892—1920).其主要工作为确定所有单的复李群.

克莱因*(Klein, Felix; 1849—1925)

克莱因生于德国杜塞尔多夫,研读于波恩大学、格丁根大学、柏林大学和巴黎大学.执教于埃尔朗根大学(1872—1875)、慕尼黑工业大学(1875—1880)、莱比锡大学(1880—1886)以及格丁根大学(1886—1913).1895年筹划卷帙浩大的《德国数学百科全书》.

科瓦列夫斯卡娅*(Kowalewska, Sofia Vassilievna; 1850—1891)

科瓦列夫斯卡娅出身于俄国贵族家庭,为能够到国外大学学习而于1868年结婚.先在海得堡大学学习数学,然后从1871至1874年在魏尔斯特拉斯处由他专门授课;从格丁根大学"缺席"(未经考试和答辩)获得博士学位,但在1884年前一直未能得到教职.1884年由米塔格-列夫勒协助获斯德哥尔摩大学教授之职.其工作领域为分析及其应用. [278]

克罗内克*(Kronecker, Leopold; 1823—1891)

克罗内克生于德国利格尼茨,就读于柏林大学、波恩大学和布雷斯劳大学;1845年获柏林大学博士学位.1845至1855年经管家庭事务,使他经济得以自立,然后返回柏林.1861年成为柏林科学院院士,并能在柏林大学开设课程.1883年继库默尔任柏林大学数学教授.他关于代数、数论和分析的论文极有深度.

库默尔*(Kummer, Ernst Eduard; 1810—1893)

库默尔生于德国索拉乌[①],就读于哈雷大学(1828—1831),然后从1832至1842年在利格尼茨预科学校任教,克罗内克就

① 今波兰扎雷.——译注

是他的学生.1842 至 1855 年为布雷斯劳大学教授,1855 年接任狄利克雷在柏林大学的教席;1861 年与魏尔斯特拉斯共同开办德国第一个纯粹数学讨论班.1855 年成为柏林科学院院士.其主要成就为"理想数"理论(用当今语言为"赋值论"),它使任何次代数数的研究成为可能.

拉格朗日 *(Lagrange,Joseph Louis;1736—1813)

拉格朗日生于意大利都灵,1755 年被任命为都灵皇家炮兵学校教授.他同一些朋友一起建立了都灵科学协会.1766 年任柏林科学院数学部主任.1787 年离柏林赴巴黎,并接受巴黎科学院的薪俸.他是度量衡委员会委员,自 1795 年天文事务所建立之日起即为该所成员.在师范学校和综合工科学校(1794 - 1799)讲授数学.在数学的各个分支中均有重要发现.

朗伯(Lambert,Johann Heinrich;1728—1777)

朗伯生于阿尔萨斯的米卢斯,12 岁时即辍学,但仍进行自学.在从事各种职业后,自 1748 至 1758 年在瑞士库尔一贵族家庭任家庭教师,其间成为巴塞尔的瑞士科学协会会员.18 世纪 60 年代初受命筹组新设的巴伐利亚科学院,但于 1764 年离慕尼黑赴柏林.1765 年任柏林科学院院士.朗伯具有百科全书式的才智,有关于哲学、逻辑学、数学和物理学的论著传世.

拉普拉斯 *(Laplace,Pierre Simon,Marquis de;1749—1827)

拉普拉斯生于法国博蒙昂诺日一农家,在该地的军事学校开始教学,并进入其数学生涯.1783 年任炮兵学校考官,1785 年 [279] 入选巴黎科学院.法国大革命期间参加筹建师范学校和综合工科学校;从法兰西研究院成立之日起即为该院院士.拿破仑很信任拉普拉斯,委以内政部长之职,但为时只有六个月.他最重要的工作是论述概率论和天体力学的论著.

瓦莱·普桑(Vallee Poussin, Charles Jean Gustave Nicolas de la; 1866—1962)

瓦莱·普桑生于比利时卢万,卒于同地;就读于卢万大学并终其一生在该大学任教,1891 年起为教授.其工作领域为分析及其应用尤其是解析数论.

勒贝格*(Lebesgue, Henri Léon; 1875—1941)

勒贝格生于法国博韦,就读于高等师范学校(1894—1897),然后在南锡公立中学任教,嗣后执教于雷恩大学(1902—1906)、普瓦蒂埃大学(1906—1910)和巴黎大学.1921 年任法兰西学院教授,次年被选为巴黎科学院院士.他在积分论、位势论和代数拓扑中引进了重要的新观念.

勒让德*(Legendre, Adrien - Marie; 1752—1833)

勒让德生于法国巴黎,在马扎林学院就读结业后即决心献身于科学研究.在巴黎军事学校讲授数学(1775—1780),1783 年入选巴黎科学院.1794 年领导公众教育委员会.1799 至 1813 年为综合工科学校考官,1813 年接替拉格朗日在天文事务所的职位.分析和数论中许多重要定理归于他名下.

莱布尼茨*(Leibniz, Gottfried Wilhelm; 1646—1716)

莱布尼茨生于德国莱比锡,就读于莱比锡大学和耶拿大学.1667 年起任迈因茨选帝侯的最高法院顾问,并出访巴黎和伦敦.1676 年不伦瑞克-吕讷堡公爵授予他汉诺威图书馆馆长之职.1700 年得勃兰登堡选帝侯之助建立柏林科学协会.在被任命为俄国彼得大帝宫廷顾问之后,他曾在维也纳居住两年(1712—1714),然后返回汉诺威,并在与世隔绝中卒于该地.作为像笛卡儿一样广博的领头学者,遗憾的是其著作尤其是逻辑

学和代数学著作长期未能出版.他引进的微积分算法和符号比牛顿的算法和符号影响大得多,并沿用至今.

李*(Lie, Marius Sophus; 1842—1899)

李生于挪威努尔菲尤尔埃德,就读于克里斯蒂安尼亚大学.毕业后以私人授课维持生计,1869—1870年冬季去柏林,邂逅克莱因,1870年夏去巴黎.1872年克里斯蒂安尼亚大学为他设立数学教席,1886年他接替克莱因在莱比锡大学的职位,1898年回克里斯蒂安尼亚大学.他创建了以他的姓氏命名的李群理论.

[280]

刘维尔*(Liouville, Joseph; 1809—1882)

刘维尔生于法国圣·奥梅尔的一个洛林家庭,就读于综合工科学校和道桥学校;任教于综合工科学校(1831—1851)、法兰西学院(1851—1879)和巴黎理学院(1857—1874).1839年成为巴黎科学院成员,1840年成为天文事务所成员.其工作与分析中许多问题有关.

李特尔伍德*(Littlewood, John Edensor; 1885—1977)

李特尔伍德生于英国罗切斯特,1903年入剑桥三一学院就读;除3年曼彻斯特大学教授外,他在剑桥度过整个大学生涯,于1928年成为剑桥大学教授直至1950年退休.他在分析和解析数论中做了深刻的工作.

罗巴切夫斯基*(Lobachevski, Nikolaï Ivanovitch; 1792—1856)

罗巴切夫斯基生于俄国下诺夫哥罗德,就读于喀山大学(1807—1812),后毕生在该校教学和工作.1827至1846年为喀山大学校长,并创建了该校天文台.

梅内克缪斯(Menaechmus;公元前4世纪中期)

梅内克缪斯可能是欧多克索斯的学生.对其生平一无所知.

梅雷(Méray,Hugues Charles Robert;1835—1911)

梅雷生于法国索恩河畔沙隆,毕业于高等师范学校,1857至1859年任教于圣康坦公立中学,然后退休7年,卜居于邻近家乡的一小村.1866年执教于里昂大学,1867年起为第戎大学教授.

梅森(Mersenne,Marin;1588—1648)

梅森生于法国曼恩的瓦泽,就读于拉弗莱谢的耶稣会学院,1611年参加小兄弟修道会.1619年后除外出旅行外住在巴黎小兄弟会的一个隐修院中.他对各种科学都有兴趣,同欧洲所有科学家进行大量通信,并在其隐修院内组建学院,自任秘书.

闵科夫斯基*(Minkowski,Hermann;1864—1909)

闵科夫斯基生于俄国阿列克索塔斯,8岁起居住于柯尼斯堡;在柯尼斯堡大学就读,内有三个学期去柏林大学.执教于波恩大学(1885—1894)、柯尼斯堡大学(1894—1896)和苏黎世综合工业大学(1896—1902).1902年格丁根大学为他设立教授席位.他在数论中发明了称为"数的几何"的方法. [281]

棣莫弗*(Moivre,Abraham de;1667—1754)

棣莫弗生于法国维特里-勒弗朗索瓦,1685年南特敕令撤销后随家庭移居英国.他在法国索莫和巴黎接受正规数学教育.在伦敦以私人讲授数学维持生计.1697年成为英国皇家学会会员.其主要工作领域为概率论.

蒙日*(Monge,Gaspard;1746--1818)

蒙日生于法国博恩,其父为商人;1765年获准入梅济耶尔皇家工程学校为技工,1766至1784年在该校讲授数学.1780年入选巴黎科学院,1783年被委任为海军军校学员的主考官.蒙日支持法国大革命,任海军部长(1792—1793),公安委员会委员,筹建综合工科学校,为拿破仑筹组埃及远征.波旁王朝复辟后被革去一切头衔和职务.其主要论著涉及微分几何和偏微分方程理论.

蒙泰尔(Montel,Paul Antoine Aristide;1876—1975)

蒙泰尔生于法国尼斯,攻读于高等师范学校,于1907年获博士学位.执教于综合工科学校和巴黎理学院(1911—1946).同时为国立高等美术学校教授,高等研究实用学校校长以及探索馆馆长;1937年成为巴黎科学院院士.其工作与分析有关.

牛顿*(Newton,Issac;1642—1727)

牛顿是遗腹子,生于英格兰沃尔索普.在剑桥三一学院受教后,于1669年继I·巴罗任剑桥大学教授.1696年离开剑桥任伦敦造币局总监.1699年成为皇家学会理事,1703年任该会主席,直至去世.他是动力学和天体力学的奠基者,又最早发明微积分的一种记号和一般算法.

[282]

诺特*(Noether,Amalie Emmy;1882—1935)

埃米·诺特生于德国埃朗根,为数学家马克斯·诺特之长女.由于当时德国妇女不得为大学正规学生,她只能在埃朗根大学和格丁根大学作为不注册的旁听生攻读数学.然而她于1904年获准注册于埃朗根大学,遂能于1907年在该校获博士学位.只是在许多有利于她的干预之后,才得以于1919年取得授课资格.1933年前在格丁根大学讲课,但从未获官方职位;1933年国

家社会党政府迫使她离校．然后她去美国避难，工作于布林莫尔学院和普林斯顿高等研究院．她的影响是创建称之为“近世代数”的数学学科的决定性因素．

欧姆(Ohm，Martin；1792—1872)

欧姆生于德国埃朗根，于1811年在埃朗根大学开始其科学生涯．在托恩预科学校任教(1817—1821)后到柏林大学，先为讲师，后任数学教授．同时任教于建筑学校(1824—1831)、联合炮兵与工程学校(1833—1852)和普通军校(1826年起)．

奥雷姆*(Oresme，Nicole；约1320—1382)

奥雷姆可能生于法国卡昂附近，就学于巴黎大学纳瓦拉学院，并在该院任教至1362年．嗣后至1377年为鲁昂大教堂牧师会牧师，1377年为利雪主教．其论著含有直到17世纪才得到开发的数学思想的萌芽．

帕斯卡*(Pascal，Blaise；1623—1662)

帕斯卡生于法国克莱蒙费朗，于1631年随其父去巴黎，1635年起参加巴黎梅森学院．1640年随其父去鲁昂，此时全家改信波尔罗亚尔隐修宗教．因病于1647年回巴黎，随即进入他的所谓“世俗”阶段，在这一期间他进行了密集的科学活动；随之又有第二次“改变信仰”．1654年后献身于富于战斗性的基督教事业，帮助詹森派信徒反对耶稣会教士．1658年后患重病．他具有非凡的数学天赋，给出了许多问题的证明．

帕施(Pasch，Moritz；1843—1930)

帕施生于德国普雷斯劳，先就学于普雷斯劳大学，然后去柏林，深受魏尔斯特拉斯和克罗内克的影响．在吉森大学度过毕生科学生涯(1870—1911)，1873年后在该校讲授严格公理形式的

[283] 几何学.

皮亚诺 * (Peano, Giuseppe; 1858—1932)

皮亚诺生于意大利斯皮内塔,十二、三岁起居于都灵,并在该市求学,继之以从事科学工作.1890 年起为都灵大学微积分学教授,1886 至 1901 年间同时任教于都灵军事学院.他是都灵科学院院士,是意大利积极参与形成 20 世纪数学概念的学派的领头人物.

庞加莱 * (Poincaré, Jules Henri; 1854—1912)

庞加莱生于法国南锡,1873 年为综合工科学校录取,在矿业学校学习工程,主要从事博士论文的准备工作.任教于卡昂大学(1879—1881)和巴黎大学(1881—1912).1887 年成为巴黎科学院院士,1908 年被选为法兰西学院院士.其创造能力和博学均可与高斯媲美,主导了他那时代数学的每个部分.

普洛克鲁斯(Proclus; 约 410—485)

普洛克鲁斯生于拜占廷,就学于亚历山大和雅典,在雅典时为柏拉图学院的成员,后领导该院,直至去世.他虽缺乏独创性,但有珍贵的“评注”传世,书中有大量从泰勒斯直至他的时代的希腊数学的详细情况.

拉马努金 * (Ramanujan, Srinivasa Aaiyangar; 1887—1920)

拉马努金生于印度邻近马德拉斯的贡伯戈纳姆,对数学的热爱使他忽视了其他课程的学习.为糊口不得不在马德拉斯港充任小官员.他与 G·H·哈代通信,告知自己的发现,遂由后者安排于 1914 年赴英国,并合作发表许多论文.由于重病不得不于 1919 年返回马德拉斯.1918 年被选为英国皇家学会会员.他在数论领域有非凡天赋,所留记事本充满各种公式,使人惊讶他

何以能提供那么多公式而又未能给出真正的证明.

黎卡提(Riccati, Jacopo Francesco; 1676—1754)

黎卡提是威尼斯贵族,在帕多瓦大学攻读法律,并对数学感到兴趣.他毕生完全献身于科学研究,拒绝了诸如圣彼得堡科学院院长这样显要的职位.他经常出任威尼斯参议院在堤坝和运河建筑事务方面的顾问.其工作涉及初期微分方程理论.

黎曼*(Riemann, Georg Friedrich Bernhard; 1826—1866)

黎曼生于德国布雷斯塞伦茨,就学于格丁根大学和柏林大学,于 1851 年在格丁根大学获博士学位,并于 1853 年取得该校 [284] 授课资格.他在格丁根大学授课,并于 1859 年接替狄利克雷的数学教授席位.肺结核缩短了他的寿命,他死于去意大利的旅途中.黎曼富有想象的天才,他的想法即使没有证明,也鼓舞了整整一个世纪的数学家.

罗伯瓦尔(Roberval, Gilles Personne de; 1602—1675)

罗伯瓦尔生于法国桑利一个地位低下的家庭,1628 年抵巴黎后,毛遂自荐进入梅森周围的科学家圈子,1634 年在竞争法兰西学院的拉米教席中获胜,遂终生担任该教席.1655 年继加桑迪任该院数学教授.1666 年巴黎科学院建立之日起即为院士.他是牛顿和莱布尼茨之前微积分的开拓者之一.

鲁菲尼(Ruffini, Paolo; 1765—1822)

鲁菲尼生于意大利瓦伦塔诺,在莫德纳大学攻读医学和数学.1788 年毕业后即被任命为教授.1796 年拿破仑军队占领莫德纳后,鲁菲尼于 1798 年拒绝宣誓效忠共和国,遂被撤销一切官方职务,但仍继续行医.拿破仑逊位后,鲁菲尼又重获应用数学和实用医学教席,直至去世.

萨凯里(Saccheri,(Giovanni)Girolamo;1667—1733)

萨凯里生于意大利圣雷莫,1685年加入耶稣会,在布雷拉耶稣会学院攻读哲学和神学.1694年被派遣外出讲授哲学,先去都灵,后去巴黎.1699年起在帕维亚大学任数学教授,直至去世.

谢林(Schering,Ernst Christian Julius;1833—1897)

谢林生于德国桑德贝根,1852年被格丁根大学录取,嗣后定居格丁根,为格丁根大学数学和天文学教授,并任高斯创建的地磁观测站主任.

施赖埃尔(Schreier,Otto;1901—1929)

施赖埃尔生于奥地利维也纳,在维也纳大学获博士学位;1926年起为汉堡大学无薪讲师,三年后因血中毒过早去世,但已在代数和群论方面做出某些引人注目的工作.

施罗德(Schröder,Friedrich Wilhelm Karl Ernst;1841—1902)

[285] 施罗德生于德国曼海姆,就学于海德堡大学和柯尼斯堡大学.1865年在苏黎世联邦综合工业大学获得讲师资格并一度任教于该校;在卡尔斯鲁厄、普福尔茨海姆和巴登-巴登任教后,于1874年任达姆斯塔德高等技术学院教授,1876年任卡尔斯鲁厄高等技术学院教授并于1890年任该院院长.其工作与逻辑和集合论有关.

施瓦茨(Schwarz,Hermann Amandus;1843—1921)

施瓦茨生于西里西亚的黑姆斯多夫,在柏林大学攻读化学和数学.1867年任哈雷大学助理教授,1869年任苏黎世联邦综合工业大学教授,1875年任格丁根大学教授.1892年接替魏尔

斯特拉斯在柏林大学的教席.他是普鲁士和巴伐利亚科学院院士,在分析中做了重要工作.

西格尔* (Siegel, Carl Ludwig; 1896—1981)

西格尔生于德国柏林,就学于柏林大学和格丁根大学;为法兰克福大学(1922—1937)和格丁根大学(1938—1940)教授.1940年借助在丹麦和挪威开的一轮会议逃脱纳粹政权统治,1940年至1951年间为普林斯顿高等研究院教授.1951年又任格丁根大学教授,1959年退休.他是20世纪数论领域最伟大的专家之一.

斯科朗(Skolem, Albert Thoralf; 1887—1963)

斯科朗生于挪威桑德斯瓦尔,就读于奥斯陆大学.在苏丹草原游历后,于格丁根大学完成学业,然后回奥斯陆大学任教(1916—1930;1938—1950).1930至1938年在卑尔根的克里斯蒂·米塞尔森学院进行独立研究.其工作涉及代数、数论和逻辑学.

史密斯(Smith, Henry John Stephen; 1826—1883)

史密斯生于爱尔兰都柏林,1840年迁居牛津,就读于牛津大学巴利奥尔学院.1860年被选为几何学教授,1874年出任牛津大学博物馆主任.1861年成为英国皇家学会会员,1877年任伦敦气象学理事会理事长.其工作涉及数论和椭圆函数.

斯蒂文* (Stevin, Simon; 1548—1620)

斯蒂文生于荷兰布鲁日,是私生子,双亲为该城富人;但对斯蒂文1583年前的情况,所知甚少.1583年他在莱顿大学注册.他作为一位“工程师”,任拿骚的毛里斯即奥伦治王子的陆海军问题顾问.他具有百科全书式的才智,有关于各种各样论题的

[286] 很多论著传世.

斯蒂费尔(Stifel, Michael; 1487—1567)

斯蒂费尔生于德国埃斯林根,1511年被任命为神父,路德的学生之一.1535年署名于维滕贝格大学,并在柯尼斯堡大学和耶拿大学讲授神学和数学.他关于代数和算术的论著引进了优于前人所用的记号,并含有对数观念的最早萌芽.

泰勒斯* (Thales; 公元前约625—约547)

泰勒斯居住于古希腊米利都,被称为古典时代在几何学中最早发明证明的学者.

韦达* (Viete, François; 1540—1603)

韦达生于法国丰特奈-勒孔特,在普瓦捷大学攻读法律.1564年为奥贝泰尔·安托瓦内特的家庭教师.1570至1573年在巴黎,受查理九世委任为雷恩的布列塔尼地方法院顾问.1580年为巴黎最高法院审查官以及皇室私人顾问.1584至1589年被宫廷放逐,后被亨利三世召回,任图尔法院推事.在法国与西班牙战争期间,韦达为亨利四世破译截获的密码信件.1602年被免职.

冯·诺伊曼(von Neumann, Johann(或 John); 1903—1957)

冯·诺伊曼生于匈牙利布加勒斯特,其父为富有的银行家.他从家庭教师接受数学训练;任教于柏林大学(1927—1929)、汉堡大学(1929—1930),1933年后执教于普林斯顿高等研究院.他参与了与战胜纳粹有关的许多科研项目,如设计原子弹.1954年成为美国原子能委员会委员.他是20世纪最深刻的分析学家之一,又是最善于把数学应用于其他科学的数学家之一.

华林(Waring, Edward; 1736—1798)

华林生于英国什鲁斯伯里,在剑桥大学马格达林学院攻读数学.1760年任剑桥大学数学教授.1763年成为英国皇家学会会员.其工作涉及代数和数论.

魏尔斯特拉斯*(Weierstrass, Karl Theodor Wilhelm; 1815—1897)

魏尔斯特拉斯生于德国奥斯登费尔特,1834年入波恩大学,四年后离校,放弃博士考试.1841年于明斯特通过国家考试,嗣后任教于几所中学.1854年获柯尼斯堡大学名誉博士,1856年被任命为柏林皇家综合工科学校教授.1856年起为柏林 [287] 大学副教授,1864年成为正教授.1856年成为柏林科学院院士.他继柯西和黎曼之后成功地把分析置于完全严格的基础之上,分析中整个一系列精致的发现应归功于他.

韦塞尔(Wessel, Caspar; 1745—1818)

韦塞尔生于挪威韦斯特贝,就读于哥本哈根大学,后为测量员和制图员.他于1798年发表的关于复数的论文一个世纪中并未为人所知.

外尔*(Weyl, Hermann; 1885—1955)

外尔生于德国埃尔姆斯霍恩,就读于格丁根大学,其中有一年在慕尼黑大学.1913年前为格丁根大学无薪讲师,1913年任苏黎世联邦综合工业大学教授.1930年接受格丁根大学的教授席位,1933年拒绝留在纳粹德国,接受了美国普林斯顿高等研究院提供的职位.他是希尔伯特的最有天才的学生,像他的老师一样,是一位全才;其最优秀的论文涉及李群,但对解析数论、泛函分析和微分几何也贡献了许多独创的思想.

扎里斯基(Zariski, Oscar; 1899—1986)

扎里斯基生于俄国科布林,就读于基辅大学、比萨大学和罗马大学;在罗马大学时为卡斯泰尔努沃的学生,于1924年获博士学位.执教于巴尔的摩的约翰·霍普金斯大学(1927—1945)、伊利诺伊大学(1946—1947)和哈佛大学(1947年后).他是只用交换代数在任意域上建立代数几何的奠基人之一.

策梅鲁*(Zermelo, Ernst Friedrich Ferdinand; 1871—1953)

策梅鲁生于德国柏林,在柏林大学、哈雷大学和弗赖堡大学攻读数学、物理学和哲学,1894年在柏林大学获博士学位.任教于格丁根大学(1899—1910)和苏黎世联邦综合工业大学(1910—1916).由于健康原因退休,居于黑森林;1916年弗赖堡大学授予他名誉教授头衔.于1935年因不满希特勒政权而离开该校,但于1946年重回.他的研究涉及力学和变分法,但其最重[288]要的工作则是关于集合论的论著.

索　引

1. 标准记号

全书使用下列标准记号.

$\mathbf{N}^*$：自然数 1，2，3，…的集合.

$\mathbf{N}=\mathbf{N}^*\cup\{0\}$：非负整数的集合.

$\mathbf{Z}$：正、负整数以及零构成的集合.

$\mathbf{Q}$：正、负有理数以及零构成的集合；其元素写为 m/n，其中 $m\in\mathbf{Z}$，$n\in\mathbf{N}^*$.

$\mathbf{R}$：实数的集合（或实直线）.

$\mathbf{R}^2$：平面.

$\mathbf{R}^3$：通常 3 维空间.

$\mathbf{C}$：复数的集合.

$x\in E$：x 属于 E，或 x 是 E 的一个元素.

$x\notin E$：$x\in E$ 的否定.

$A\subset E$：A 是 E 的子集，或 A 包含于 E 中.

$A\not\subset E$：$A\subset E$ 的否定.

(x,y)：对象的偶.

$E\times F$：偶 (x,y) 的集合，其中 $x\in E$，$y\in F$.

$E_1\times E_2\times\cdots\times E_n$：$n$ 元组 $(x_1,x_2,\cdots,x_n)$ 的集合，其中 $x_1\in E_1$，$x_2\in E_2$，…，$x_n\in E_n$.

E^n：所有 E_j 都等于 E 时的 $E_1\times E_2\times\cdots\times E_n$.

$\mathfrak{P}(E)$：E 的子集的集合.

$\varnothing$：空集.

$f:E\to F$, $E\xrightarrow{f}F$, $x\longmapsto f(x)$:集合 E 到集合 F 中的映射 f.

$a\equiv b(\mathrm{mod}\,m)$:模整数 m 的同余式,即 $a-b$ 是 m 的整数倍.

$\mathrm{d}y/\mathrm{d}x$, $y'(x)$:实变量的实值函数 $x\longmapsto y(x)$的导数.

$\int_a^b f(x)\mathrm{d}x$:实函数 $f(x)$在区间$[a,b]$上的积分.

$\mathfrak{S}_n$:n 个对象的排列构成的对称群,其阶为 $n!=1\cdot2\cdot3\cdots\cdots n$.

2. 专名索引[①]

① 译者对原书索引作了增补和修改,并对正文进行了相应的处理,使得所有专名出现其定义或说明时都用另外的字体标志.

索引中第一个罗马数字表示出现定义时专名所在的章次,第二个数字表示节次,下类推."Ap."表示附录.——译注

五画

六　画

七 画

八 画

九 画

十 画

十 一 画

十 二 画

3. 人名索引[①]

① 本索引系译者所编,收入了正文中出现的全部人名,并指明了出现的章节.索引中第一个罗马数字表示出现的章次,第二个数字表示节次,下类推."Ap."表示附录,"Int."表示导言.人名后附有＊号者附录中有其小传.——译注

五 画

六 画

八 画

九 画

十 画

十 一 画

十 二 画